

3642	AAAATGTGA	ACTGCAC	TCCCGCTG	ACTG	CCAG	AGAAAAA	TATATG	AGAC	CCCTTGA	AGACT	3701																		
QY	301	CCAG	AACTTCA	AGAGCC	CCAG	ATGAG	CTG	AG	CTTCA	AGCTTG	GGCC	CAAGT	GGT	AT	360														
Db	3702	CCAG	AACTTCA	AGAGCC	CCAG	ATGAG	CTG	AG	CTTCA	AGCTTG	GGCC	CAAGT	GGT	AT	3761														
QY	361	CAA	GGAGT	CCCTG	GCAG	CCCGT	GGG	CGAT	CTG	CTC	ATT	CA	CTC	CTC	CA	AGAT	CA	CTC	CA	420									
Db	3762	CAA	GGAGT	CCCTG	GCAG	CCCGT	GGG	CGAT	CTG	CTC	ATT	CA	CTC	CTC	CA	AGAT	CA	CTC	CA	3821									
QY	421	GAA	AGTCA	AGG	CAC	TTT	CG	AG	GA	AAAT	TG	CC	CT	CT	TGA	AA	GAG	AA	CGT	GA	480								
Db	3822	GAA	AGTCA	AGG	CAC	TTT	CG	AG	GA	AAAT	TG	CC	CT	CT	TGA	AA	GAG	AA	CGT	GA	3881								
QY	481	TG	ACCTT	GC	TGC	CA	CTT	CA	CA	CTT	TG	GG	GA	AT	T	CA	AG	CT	CA	CG	TA	CA	AG	CA	540				
Db	3882	TG	ACCTT	GC	TGC	CA	CTT	CA	CA	CTT	TG	GG	GA	AT	T	CA	AG	CT	CA	CG	TA	CA	AG	CA	3941				
QY	541	TCT	GAA	AG	CTT	CA	CA	CA	AG	AT	TG	GA	AG	CTT	CT	G	AG	GT	GC	CG	CG	TA	CA	AG	CA	600			
Db	3942	TCT	GAA	AG	CTT	CA	CA	CA	AG	AT	TG	GA	AG	CTT	CT	G	AG	GT	GC	CG	CG	TA	CA	AG	CA	4001			
QY	601	CG	AG	CTG	CA	TGA	AG	CC	CA	AG	GA	CTT	TG	GT	CC	CA	AG	AT	CT	CA	AG	CA	CTT	CTT	CA	AG	CTC	660	
Db	4002	CG	AG	CTG	CA	TGA	AG	CC	CA	AG	GA	CTT	TG	GT	CC	CA	AG	AT	CT	CA	AG	CA	CTT	CTT	CA	AG	CTC	4061	
QY	661	TG	TC	CAG	GGT	CC	CT	GC	GC	AG	AG	AG	CA	CT	TC	CC	CA	AA	CA	AG	TC	CT	ACT	AT	TA	CA	CA	720	
Db	4062	TG	TC	CAG	GGT	CC	CT	GC	GC	AG	AG	AG	CA	CT	TC	CC	CA	AA	CA	AG	TC	CT	ACT	AT	TA	CA	CA	4121	
QY	721	CG	AG	CTCA	AA	CA	ACTT	GC	TG	GG	AC	AT	CC	CA	AA	AT	G	CA	AG	CT	CT	CA	AG	CTT	TC	780			
Db	4122	CG	AG	CTCA	AA	CA	ACTT	GC	TG	GG	AC	AT	CC	CA	AA	AT	G	CA	AG	CT	CT	CA	AG	CTT	TC	4181			
QY	781	TG	AC	CTG	AA	TAT	GT	CA	AT	CT	CA	AG	CT	TA	TAG	AG	CT	GC	CA	TG	AA	CA	CTC	CG	AA	CA	CTG	CA	840
Db	4182	TG	AC	CTG	AA	TAT	GT	CA	AT	CT	CA	AG	CT	TA	TAG	AG	CT	GC	CA	TG	AA	CA	CTC	CG	AA	CA	CTG	CA	4241
QY	841	GAA	GCC	CTT	TG	CT	TG	AT	CT	CT	TG	AG	CT	GT	CA	CT	GT	CA	CT	AT	GT	AG	CA	CA	CA	CA	900		
Db	4242	GAA	GCC	CTT	TG	CT	TG	AT	CT	CT	TG	AG	CT	GT	CA	CT	GT	CA	CT	AT	GT	AG	CA	CA	CA	CA	4301		
QY	901	CA	ACT	CTCA	AG	AAAA	TGA	CC	AG	CC	AT	TG	GA	TAT	CT	CG	AA	GA	TAT	TA	T	GT	TGA	CC	AC	960			
Db	4302	CA	ACT	CTCA	AG	AAAA	TGA	CC	AG	CC	AT	TG	GA	TAT	CT	CG	AA	GA	TAT	TA	T	GT	TGA	CC	AC	4361			
QY	961	TAT	TAT	TAG	CCG	CT	GC	AG	CA	AG	CA	CA	AA	TT	TG	CT	CA	AG	CT	CC	CT	CT	CG	TG	GA	1020			
Db	4362	TAT	TAT	TAG	CCG	CT	GC	AG	CA	AG	CA	CA	AA																

```

: Sequence 60: Application US/09484970B
: Patent No. 6426186
: GENERAL INFORMATION:
: APPLICANT: Jones, Karen A.
: APPLICANT: Volkmuth, Wayne
: APPLICANT: Walker, Michael G.
: TITLE OF INVENTION: BONE REMODELLING GENES
: FILE REFERENCE: PB-0014 US
: CURRENT APPLICATION NUMBER: US/09/484,970B
: CURRENT FILING DATE: 2000-01-18
: NUMBER OF SEQ ID NOS: 172
: SOFTWARE: PERL Program
: SEQ ID NO 60
: LENGTH: 13977
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: incyle ID No. 6426186 229357.11CBI
: NAME/KEY: unsure
: LOCATION: 11721-11761, 12294, 13969
: OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match      99.28; Score 1290; DB 4; Length 13977;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY      1 CGACTTCCGACGATTGAGAGGAGAACGATGTACATAGGGCCCTCAAGAGGGAATTGAA 60
DB      8716 CGACTTCCGACGATTGAGAGGAGAACGATGTACATAGGGCCCTCAAGAGGGAATTGAA 8775

OY      61 AACTAAGAACTGTATATCATGATGACTCTTGAGACTGTACGAATATTCTGACAGCA 120
DB      8776 AACTAAGAACTGTATATCATGATGACTCTTGAGACTGTACGAATATTCTGACAGCA 8835

OY      121 GCCTTTGAGAGCATGAGAAACTCTTACAGAGGCCAGAGAGCTGCTCTGAGAGAG 180
DB      8836 GCCTTTGAGAGCATGAGAAACTCTTACAGAGGCCAGAGAGCTGCTCTGAGAGAG 8895

OY      181 AGCCCAAGATGTCCTGCGCTTACGAAAGCAGGCTGAGAGGTCATATCTAGTGGGA 240
DB      8896 AGCCCAAGATGTCCTGCGCTTACGAAAGCAGGCTGAGAGGTCATATCTAGTGGGA 8955

OY      241 AAAATTGAACCTGCACTCCGCTGACTGCGAGAGAAATATGATGAGACCTTGAAAGACT 300
DB      8956 AAAATTGAACCTGCACTCCGCTGACTGCGAGAGAAATATGATGAGACCTTGAAAGACT 9015

OY      301 CCAGGAACCTCAAGAGGCGCAGGATGAGTGTGACCTCAACCTCGCCCAAGCTGAGTGAT 360
DB      9016 CCAGGAACCTCAAGAGGCGCAGGATGAGTGTGACCTCAACCTCGCCCAAGCTGAGTGAT 9075

OY      361 CAAAGGATCCTGCGAGCCGCTGGCGCATCTTCTTCATTGACTCTCTCCAAGATCACTTCGA 420
DB      9076 CAAAGGATCCTGCGAGCCGCTGGCGCATCTTCTTCATTGACTCTCTCCAAGATCACTTCGA 9135

OY      421 GAAAGTGAAGGCACTTGGAGAGAAATTTGGCCCTCTGAAGAGAAAGCTGAGCAACGTCAA 480
DB      9136 GAAAGTGAAGGCACTTGGAGAGAAATTTGGCCCTCTGAAGAGAAAGCTGAGCAACGTCAA 9195

OY      481 TGACCTTGCTGCGCAGCTTACCACTTTGGCAATTCAAGCTCTCAACGTAATACCTCAGAC 540
DB      9196 TGACCTTGCTGCGCAGCTTACCACTTTGGCAATTCAAGCTCTCAACGTAATACCTCAGAC 9255

OY      541 TCTGGAAGACCTGAAACCAAGATGGAAGCTTTCGAGAGTGGCCGTGAGAGCAGAGTCAG 600
DB      9256 TCTGGAAGACCTGAAACCAAGATGGAAGCTTTCGAGAGTGGCCGTGAGAGCAGAGTCAG 9315

OY      601 GCAAGCTGATGAAGCCCAAGGAGCTTTGGTCAAGACCTCAAGCACTTTCTTTCCAGTC 660
DB      9316 GCAAGCTGATGAAGCCCAAGGAGCTTTGGTCAAGACCTCAAGCACTTTCTTTCCAGTC 9375

OY      661 TGTCACAGGTCCTCGGAGAGAGCACTCGCCCAAAAGTAGTGCCTCAATATATCAACCA 720

```


DB 9376 TGTCCAGGCTCCTGGAGAGAGCCATCTTGCCAAAGAGCCCTACTATATACCA 9435
QY 721 CGAGACTCAACCACTTGTGTGGAGCCATCCAAATAGACAGCTCTACAGCTTTAGC 780
DB 9436 CGAGACTCAACCACTTGTGTGGAGCCATCCAAATAGACAGCTCTACAGCTTTAGC 9495
QY 781 TGAACCTGAATAGTCAAGTTCACACTTATAGAGATGCGCATGAAACTCCGAAGCTGA 840
DB 9496 TGACCTGAATAGTCAAGTTCACACTTATAGAGATGCGCATGAAACTCCGAAGCTGA 9555
QY 841 GAAGCCCTTGTGTGGATCTCTTGAAGCTGTGAGCTGTGATGATGCTTGAAGCAGA 900
DB 9556 GAAGCCCTTGTGTGGATCTCTTGAAGCTGTGAGCTGTGATGATGCTTGAAGCAGA 9615
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGATATCTGACATTAATTAATTTGTTAGCAC 960
DB 9616 CAACCTCAAGCAAAATGACCAAGCCCATGATATCTGACATTAATTAATTTGTTAGCAC 9675
QY 961 TATTATAGACCGCTGTGAGAGAGACCAAAATTTGTGCAAGCTCTCTGTGCTGGA 1020
DB 9676 TATTATAGACCGCTGTGAGAGAGACCAAAATTTGTGCAAGCTCTCTGTGCTGGA 9735
QY 1021 TATGTCTGAACCTGTGCTGTGATTTATGATAGAGAGACCAAGAGAGAGCTGT 1080
DB 9736 TATGTCTGAACCTGTGCTGTGATTTATGATAGAGAGACCAAGAGAGAGCTGT 9795
QY 1081 CCTGTCTTTAAACTGTGACATCTTCCCTGTGTAAAGACATTTGGAAACAGATACAG 1140
DB 9796 CCTGTCTTTAAACTGTGACATCTTCCCTGTGTAAAGACATTTGGAAACAGATACAG 9855
QY 1141 ATACCTTTCAAGCAAGTGTGACATCTTCCCTGTGTAAAGACATTTGGAAACAGATACAG 1199
DB 9856 ATACCTTTCAAGCAAGTGTGACATCTTCCCTGTGTAAAGACATTTGGAAACAGATACAG 9915
QY 1200 TCCCTGTGATGATTTCTATCCAAATTCAGAGACATTTGGATGATGCTTTGGAG 1259
DB 9916 TCCCTGTGATGATTTCTATCCAAATTCAGAGACATTTGGATGATGCTTTGGAG 9975
QY 1260 GCGATACATGTAGCCCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
DB 9976 GCGATACATGTAGCCCAAGTGTCCGAGCTGCTTCCAAATTTG 10017

RESULT 3

US-08-836-022A-10/c
Sequence 10, Application US/08836022A
Patent No. 600157
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
APPLICANT: Wilson, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Chen, Shu-Jen
APPLICANT: Weltman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,022A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Maty E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVN. 008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-836-022A-10
Query Match 87.3%; Score 1135.2; DB 3; Length 19307;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 2 GACTTTCACAGCTTCAGAGACAGACGATGATAGAGCTTCAAGAGGAATTGAA 61
DB 5977 GATTCCAGAGTTGAGAGAGAGATGATATATAGAGCTTCAAGAGGAATTGAA 5918
QY 62 ACTAAGAACTCTGATATGATGATCTTTGAGACTGTACGATATTTCTGACAGAGCAG 121
DB 5917 ACTAAGAACTCTGATATGATGATCTTTGAGACTGTACGATATTTCTGACAGAGCAG 5858
QY 122 CCTTGGAGAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTGAGAGAGA 181
DB 5857 CCTTGGAGAGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTGAGAGAGA 5798
QY 182 GCCAGATGTCACTGCTGCTTCTACAGAAAGAGCTGAGAGCTCAATGATGAGGAA 241
DB 5797 GCTCAGATGTCACTGCTGCTTCTACAGAAAGAGCTGAGAGCTCAATGATGAGGAA 5738
QY 242 AATTTGAACCTGACCTCCGCTGCTGAGAGAGAAATATGATAGACCCTTGAAGACATC 301
DB 5737 AATTTGAACCTGACCTCCGCTGCTGAGAGAGAAATATGATAGACCCTTGAAGACATC 5678
QY 302 CAGGAATTCAGAGAGCCAGATGAGCTGACCTGACCTGAGCCAGAGCTGAGTATC 361
DB 5677 CAGGAATTCAGAGAGCTGAGCTGAGCTGACCTGACCTGAGCCAGAGCTGAGTATC 5618
QY 362 AAGGATCTCGGACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 421
DB 5617 AAGGATCTCGGACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 5558
QY 422 AAGTCAAGGCACTTGAGAGAGAAATGAGCTCTGAAAGAGAGAGGAGCCAGTCAAT 481
DB 5557 AAGTCAAGGCACTTGAGAGAGAAATGAGCTCTGAAAGAGAGAGGAGCCAGTCAAT 5498
QY 482 GACCTTGTGAGCACTTGAGAGAGAAATGAGCTCTGAAAGAGAGAGGAGCCAGTCAAT 541
DB 5497 GACCTTGTGAGCACTTGAGAGAGAAATGAGCTCTGAAAGAGAGAGGAGCCAGTCAAT 5438
QY 542 CTGGAAGACTGAAACAGATGAGAGCTTTGCAAGTGGCCCTGAGAGCCAGAGTCAAG 601
DB 5437 TTGGAAGATGTGAATCCAGATGAGAGCTTTGCAAGTGGCCCTGAGAGCCAGAGTCAAG 5378
QY 602 CAGCTGATGAAGCCAGAGAGCTTTGTCAGACATCTCAGACATTTTTCACAGTCT 661
DB 5377 CAGCTGATGAAGCCAGAGAGCTTTGTCAGACATCTCAGACATTTTTCACAGTCT 5318
QY 662 GTCCAGGCTCTCTGGAGAGAGAGCTTGTGCCAAAGAGTGGCTTACTATATCAACAC 721
DB 5317 GTCCAGGCTCTCTGGAGAGAGAGCTTGTGCCAAAGAGTGGCTTACTATATCAACAC 5258
QY 722 GAGACTCAACCACTTGTGAGAGCACTTCCAAATATGAGAGAGCTTACAGCTTTAGCT 781
DB 5257 GAGACTCAACCACTTGTGAGAGCACTTCCAAATATGAGAGAGCTTACAGCTTTAGCT 5198

QY 782 GACCTGATAATGTCAGATTCCTAGCTTATAGAGCTGCGATGAAACTCCAGAGCTGAC 841
Db 5197 GACCTGATAATGTCAGATTCCTAGCTTATAGAGCTGCGATGAAACTCCAGAGCTGAC 5138
QY 842 AAGGCCCTTGTGCTGATCTCTTGAAGCTGCTGACCTGATGATGCTTGGACACGAC 901
Db 5137 AAGGCCCTTGTGCTGATCTCTTGAAGCTGCTGACCTGATGATGCTTGGACACGAC 5078
QY 902 AACCTGACAAATGACCAAGCCATGATATCTGACGATTTATTAATGTTGACCACT 961
Db 5077 AACCTGACAAATGACCAAGCCATGATATCTGACGATTTATTAATGTTGACCACT 5018
QY 962 ATTATGACCCGCTGAGCAAGACCAAAATTTGCTCAAGCTCTCTCTGCTGAT 1021
Db 5017 ATTATGACCCGCTGAGCAAGACCAAAATTTGCTCAAGCTCTCTCTGCTGAT 4958
QY 1022 ATGCTGCTGACCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1081
Db 4957 ATGCTGCTGACCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 4898
QY 1082 CTGCTTTTAAATGCGATCATTTCCCTGCTGTAAGCACATTTGGAAGCAAGTACAGA 1141
Db 4897 CTGCTTTTAAATGCGATCATTTCCCTGCTGTAAGCACATTTGGAAGCAAGTACAGA 4838
QY 1142 TACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACGCGAGCTGGGCTC 1201
Db 4837 TACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACGCGAGCTGGGCTC 4778
QY 1202 CTCTGATGATTTCTATCCAAATTCACAGAGTGGTGAAGTGGATCCTTTGGGGGC 1261
Db 4777 CTCTGATGATTTCTATCCAAATTCACAGAGTGGTGAAGTGGATCCTTTGGGGGC 4718
QY 1262 AGTAACTGAGCCAGTGTCCGAGCTGCTTCAATTTG 1301
Db 4717 AGTAACTGAGCCAGTGTCCGAGCTGCTTCAATTTG 4678

RESULT 4
US-09-427-048A-10/c
Sequence 10, Application US/09427048A
Patent No. 6203975
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
Wilson, James M.
Fisher, Krishna J.
Chen, Shu-Jen
Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
Methods of Use Thereof
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E
REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: GNPVN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10
Query Match 87.3%; Score 1135.2; DB 3; Length 19307;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 2 GACTTTCAGAGCTTCAGAGCAGATGATGATGATGATGATGATGATGATGATGAT 61
Db 5977 GACTTTCAGAGCTTCAGAGCAGATGATGATGATGATGATGATGATGATGATGAT 5918
QY 62 ACTAAGAACTGTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121
Db 5917 ACTAAGAACTGTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5858
QY 122 CTTTGGAAAGGACTGAGAACTTACCAAGAGCCCAAGAGCTGCTCTGAGAGAGA 181
Db 5857 CTTTGGAAAGGACTGAGAACTTACCAAGAGCCCAAGAGCTGCTCTGAGAGAGA 5798
QY 182 GCCAGATGTCACCTGCTGCTTCTCAAGAGAGCTGAGAGAGTCAATATGAGTGGAAA 241
Db 5797 GCTCAGATGTCACCTGCTGCTTCTCAAGAGAGCTGAGAGAGTCAATATGAGTGGAAA 5738
QY 242 AATTGAACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 301
Db 5737 AATTGAACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5678
QY 302 CAGGAATCTCAGAGCCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 361
Db 5677 CAGGAATCTCAGAGCCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 5618
QY 362 AAGGATCTCTGAGCCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
Db 5617 AAGGATCTCTGAGCCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5558
QY 422 AAGTCAAGGCACTTGCAGAGAAATTTGCCCTCTGAAAAGAGAGAGTGAAGCCATCAT 481
Db 5557 AAGTCAAGGCACTTGCAGAGAAATTTGCCCTCTGAAAAGAGAGAGTGAAGCCATCAT 5498
QY 482 GACCTTGTGCTGCGAGCTTACCACTTTGGGCACTTCACTGCTCAAGCTATACCTGACACT 541
Db 5497 GACCTTGTGCTGCGAGCTTACCACTTTGGGCACTTCACTGCTCAAGCTATACCTGACACT 5438
QY 542 CTGGAAGACCTGAGACCAAGATGAGAGCTTCTGAGGCTGCTGAGAGCCGAGCTGAG 601
Db 5437 TTGGAAGACCTGAGACCAAGATGAGAGCTTCTGAGGCTGCTGAGAGCCGAGCTGAG 5378
QY 602 CAGCTGATGAGAGCCCAAGGAGCTTTGGTCCAGATCTGAGCACTTCTTCCAGGCT 661
Db 5377 CAGCTGATGAGAGCCCAAGGAGCTTTGGTCCAGATCTGAGCACTTCTTCCAGGCT 5318
QY 662 GTCCAGAGGCTCTGAGAGAGAGCCATCTGAGCCAAAGAGTGCCTTACTATATCAACAC 721
Db 5317 GTCCAGAGGCTCTGAGAGAGAGCCATCTGAGCCAAAGAGTGCCTTACTATATCAACAC 5258
QY 722 GAGACTCAAAACATTTGCTGAGACCATCTCCAAAGAGAGAGCTTACAGCTTTAGCT 781
Db 5257 GAGACTCAAAACATTTGCTGAGACCATCTCCAAAGAGAGAGCTTACAGCTTTAGCT 5198
QY 782 GACCTGATAATGTCAGATTCCTAGCTTATAGAGCTGCGATGAAACTCCAGAGCTGAC 841
Db 5197 GACCTGATAATGTCAGATTCCTAGCTTATAGAGCTGCGATGAAACTCCAGAGCTGAC 5138

OY	842	AAGGCCCTTGGTTGGAATCTCTTGACCGCTGTACGTGCATGTAATGATCCCTTGACAACAGC	901
Dd	5137	AAGGCCCTTGGTTGGAATCTCTTGACCGCTGTACGTGCATGTAATGATCCCTTGACAACAGC	5078
OY	902	AACCTCAAGCAAATGACCAGCCCCATGATATCTGCGAATTATTAATTGTTTGACAACCT	961
Dd	5077	AACCTCAAGCAAATGACCAGCCCCATGATATCTGCGAATAATTAATTCGTTTGACTCA	5018
OY	962	ATTATGACCGGCTGTGAGACAAGACAACAATTTGGTCAATCGTCCCCTCTGCGGTGAT	10211
Dd	5017	ATTATGATCGCTGTGAGACAAGACAACAATCTGTGTCAATGTCCCTCTCTGTGTGAT	4958
OY	1022	ATGTGCTGAACTGGCTGTGTAATGTTATGATACGGAACGAAACGAGGAGATCCGTGTC	10811
Dd	4957	ATGTGCTCAACTGGCTTCTCATGTTATGATACGGAACGAAACGAGGAGATCCGTGTC	4898
OY	1082	CTGTCTTTTMAAACCTGGCATCATTTTCCCTGTGTAAAGACATTTTGAAGAACAATGTAACA	11411
Dd	4897	CTGTCTTTTMAAACCTGGCATCATTTTCTGTGTAAAGACACTTTGGAAGAAACAATGTAACA	4838
OY	1142	TACCTTTTCAAGCAGTAGTGCAAGTTCACACAGATTGTCACAGCGCAGGCTGGGCTTC	12011
Dd	4837	TACCTTTTCAAGCAGTAGTGCAAGTTCACACAGCTGTTTGTGACAGCGCTAAGCTGGGCTTT	4778
OY	1202	CTTCTGCATGATTTATCCAAATTCCAAGACAGATTGGGTGAAGTTGCATCTTTGGGGGC	12611
Dd	4777	CTTCTGCATGATTTATTCACAAATCCCAAGACAGATTGGGGTGAAGTTGCTCTTTGGGGGC	4718
OY	1262	AGTACATTGAGCCAGGTGTCGGAGCTGCTTCCAAATTG	1301
Dd	4717	AGTACATTGAGCCAGGTGTCAGGAGCTGCTTCCAAATTG	4678

RESULT 5
US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413

```

? APPLICANT: Tinsley, Jonathan M
? APPLICANT: Davies, Kay E
? TITLE OF INVENTION: Utrrophin gene expression
? FILE REFERENCE: 620-42
? CURRENT APPLICATION NUMBER: US/09/091,501B
? CURRENT FILING DATE: 1998-06-18
? PRIOR APPLICATION NUMBER: PCT/GB96/03156
? PRIOR FILING DATE: 1996-12-19
? PRIOR APPLICATION NUMBER: GB 9525962.8
? PRIOR FILING DATE: 1995-12-19
? PRIOR APPLICATION NUMBER: GB 9615797.9
? PRIOR FILING DATE: 1996-07-26
? PRIOR APPLICATION NUMBER: GB 9622174.2
? PRIOR FILING DATE: 1995-10-24
? NUMBER OF SEQ ID NOS: 15
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 7
? LENGTH: 6045
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (11)..(6037)
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Chimera
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (724f..(758)
? OTHER INFORMATION: Precise residue is left open
? US-09-091-501B-7

```

Query Match	42.6%	Score 553.8	DB 4	Length 6045
Best Local Similarity	64.7%	Pred. No. 3.8e-169		
Matches 844	Conservative 0	Mismatches 452	Indels 9	Gaps 1

OY	2	GACCTTCAGAGCTCCAGAAAGCAACAAAGATACATAGAGCCCTTCAGAGGAAATTGAAA	61
Dp	3506	GATGTTCCAGCCTTACAGCCTCAAGTATGACCATTTGTAAGGCCCTGAGACGGGAGTTAAAG	3565
OY	62	ACTTAAAGACCTGTAAATCATGAGTACTTGGACTGTAGACTGTACGAATATTTCTACACAGAGAG	121
Dp	3566	GAGAAAGAAATATTTGTCTCCGAAATGCTGTGACACAGGCCCAAGTTTCTTGGCTGATACG	3625
OY	122	CCCTTGGAAAGACTAGAGAAACT-----CTACACAGAGGCCAGAGAGCTGCCTCT	172
Dp	3566	CCAATTAGAGCCCTGTAAGAGCCAAAGAAACCTACATCAATCAAAAAACGAATTAACCTCT	3685
OY	173	GAGGAGAGAGCCCAAGAAATGTCACTCGGCTTTACGAAAGCAGGCTGAGAGAGTCAATACT	232
Dp	3566	GAGGAGAGAGCCCAAGAAATGTCCAAAGCCATGGCAACAGCTTCTGAAAGTCAAGAA	3745
OY	233	GAGTGGAGAAAAATTGAACCTGCACCTCCGCTGACTGCGAGAGAAAAAATAGATGAGACCTT	292
Dp	3746	AAATGGGAAAGTCAAAATGCTGTATCATAGCAATTTGGCAAAAGCAAGTGAACAAGCAATTG	3805
OY	293	GAAGAAGCTCCAGAACTTCAAGAGGCCAGATGACTGTGACCTCAAGCTGGCCAACT	352
Dp	3806	GAGAAACTCGAGACCTGCAAGGAGCATATGATGATCACTGACCGCTGACATAGAGAGCA	3865
OY	353	GAGGTGATCAAGGAGATCCTGGCAGCCCGTGGGCATCTCCATTTGACTCTCTCCAAAT	412
Dp	3866	GAGTCCGTGGGAAATGACTGGAAACCCGTGGGAACATTAATCACTTGAATCTGCTGAGAT	3925
OY	413	CACCTCGAGAAAGTCAGAGCACTTCAGAGAGAAATTTGGCCTCTGAAAGAGACGTGAGC	472
Dp	3926	CACATTTAAAAAATCAATGAGCATTTAGAGAAAAATTCACACCAATCAACTTTAAAGTTAA	3985
OY	473	CAGCTCATGACCTTGTCCGCAAGCTTACCACTTTGGGACATTGAGCTCTCACGATTAAC	532
Dp	3986	ACGATGATGATTTATTCAGTCAAGCTGTCTCCACTTGACCTGCATCCCTCTTAAAGATG	4045
OY	533	CTCAGCACTCTGGAAGACCTGGAACACACAGATGGAAGCTTCTGACAGTGGCCGTGAGAGC	592
Dp	4046	TTTCGCCAGCTAGATGACCTTTAAATGTCGATGGAACCTTTTACAGGTTTCTGTGATGAT	4105
OY	593	CGAGTCAGAGAGCTGCATGAGACCCACAGGGAACCTTGTGTCAGAGATCTCACACATTTCTT	652
Dp	4106	CGCCTTAAACAGCTTCAGGAAGCCCAACAGATTTTGGACCATCTCTCAGACATTTTCTC	4165
OY	653	TCCACGCTGTCTCCAGGGTCCCTGGGAGAGACCATCTGCGCCAAACAAAGTGCCTTACTAT	712
Dp	4166	TCTACGTCAGTCCAGCTGCGGTGGCAAAAGATTCATTTTCACTAATTAAGTGCCTTATAC	4225
OY	713	ATCAACACAGAGATCTCAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTCACAG	772
Dp	4226	ATCAACATCAAAACAGACCAACCACTGTGTGGAGCCATCTTAAATGACCGAATCTTTCAA	4285
OY	773	TCTTTAGCTGACCTGAATTAATGTCAKATCTCAGCTTATAGGACTGCACTGAACCTCCGA	832
Dp	4286	TCCCTTCTGACCTGAATTAATGTACGTTTTTCTGCCATCCGTAACAGCAATTAATATCCGA	4345
OY	833	AGACTGACAGAGCCCTTGTCTTGGATCTCTTGAAGCTGTGACGCTGACGTGATGTGACCTTG	892
Dp	4346	AGACTTAAAAAAGACATATGTTTGAATCTCTTAAAGTTAGATCAACAAATGAATTTTC	4405
OY	893	GACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGATATCTGCAATTAATTAATGT	952
Dp	4406	AAAAGAGACAAAGTTGAACAAAAGACACAGCTCTCTCAGTGTTCAGAGTGTATCAACGT	4465
OY	953	TTGACCACTATTTTGAACGCTGAGACAGAGACAACAATTTGTCAACGTCCTCTC	1012
Dp	4466	CTGACAAACAATTATGATGACTTGAGCAATATGATTAAGGACCTGTCTCAACTC	4525
OY	1013	TGCGTGAATATGTCTGGAAGCTGGCTGTGAATTTTATGATAGGAGCAACAGGAGG	1072
Dp	4526	TGTGTTAATATGTCTCAATGCTGTGTCTCATGTCTATGACAGGGTCTGAACTGGAAAA	4585

QY 1073 ATCCGTCCTGCTTTTAAACTGSCATATTCCTGTGTAAGCAATTGGAGAC 1132
DB 4586 ATTAGAGTGCAGAGCTGAAAGATGATGATGCTCTCCAAAGGCTCTTGGAGAA 4645
QY 1133 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACGACG 1192
DB 4646 AATATCAGATATCTCTTTAAGAGATGCGGGCCGACAAATGCTGACGAGAGCAG 4705
QY 1193 CTGGGCTCTCTTGCATGATTTATCCAAATTCAGACAGTTGGGTGAAGTGCATCC 1252
DB 4706 CTGGGCTCTTACTTCAATGATGATCCATCCAGATCCCGGACAGTGAAGTGAAGCT 4765
QY 1253 TTGGGGGAGTAACTATGAGCCAGTGTCCGGAGCTGTCCAA 1297
DB 4766 TTGGAGGAGTAAATATGAGCTGATGCTGTCCAA 4810

RESULT 6
US-09-091-501B-9
Sequence 9, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathan M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091, 501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 10320
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(10312)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Full length
FEATURE:
NAME/KEY: misc_feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match 42.6%; Score 553.8; DB 4; Length 10320;
Best Local Similarity 64.7%; Pred. No. 5.3e-169;
Matches 844; Conservative 0; Mismatches 452; Indels 9; Gaps 1;

QY 2 GACCTTCAGAGCTTCAAGACAGACATGATACATGAGGCTTCAAGAGGCAATTGAA 61
DB 7781 GATGTCAGACCTTACAGCTCAAGATACATGATGAGGCTTCAAGAGGCAATTGAA 7840
QY 62 ACTAAAGAACTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121
DB 7841 GAGAAAGAAATTTCTGCTCTGATGATGATGATGATGATGATGATGATGATGAT 7900
QY 122 CTTTGAAGAGCTAGAAACT-----CTACGAGAGCCAGAGAGCTGCTCT 172
DB 7901 CCAATTGAGGCTTGAAGAGCCAGAAAGAACTTCAATCAAAAGAGATTAATCTCT 7960
QY 173 GAGGAGAGAGCCAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 232
DB 7961 GAGGAGAGAGCCAGAAAGAAATGATGATGATGATGATGATGATGATGATGATGAT 8020

QY 233 GAGTGGGAAATATGAACTGCACTCCGCTGACTGCGAGAGGAAATATGATGAGACCTT 292
DB 8021 AATGGGAAATCTAAATATGCTGTATGATGATGATGATGATGATGATGATGATGATG 8080
QY 233 GAAAGACTCCAGGAACCTTCAAGAGCCAGATGAGTGGACCTTCAAGCTGCGCAAGCT 352
DB 8081 GAGAACTCAAGAGACTTGCAGAGGCTTATGATGATGATGATGATGATGATGATGATGAT 8140
QY 353 GAGGTGATCAAGGATCTGAGAGCCGCTGAGGCTGATCTCTCATTTGACTCTTCAAGAT 412
DB 8141 GAGTCCGAGAAATGCTGAGAGCCGCTGAGGATCTTACTGATGATGATGATGATGAT 8200
QY 413 CACCTCGAGAAAGTCAAGGACTTCAAGAGAAATGAGGCTTCAAGAGGAGAGAGT 472
DB 8201 CACATTTGAAATATATGAGCAATTGAGAGAAATGAGCAATCAATTTAAAGTTAA 8260
QY 473 CACGTCAATGACCTTGTCTGCGCAGCTTACCACTTTGAGCAATTCAGCTCTCAACG 532
DB 8261 ACCGTGATGATTTATCCAGTCAAGTGTCTCAGCTTGAACCTGCAATCCCTCTTAAAG 8320
QY 533 CTGAGCACTCTGGAAGACTTGAACCAAGATGAGAGCTTCTGAGAGTGGCCGTGAGAG 592
DB 8321 TCTGCCAGCTAGAGACTTAAATGAGATGAACTTTACAGGTTTCTGTGATGAT 8380
QY 593 CGAGTCAGAGCTGATGAGAGCCCAAGGACTTTGTCAGAGATCTCAGACCTTCTT 652
DB 8381 CGCCTTAAAGACTTCAAGAGCCCAAGAGATTTTGAACCATCTCTCAGACATTTCTC 8440
QY 653 TCCAGCTGTGTCAGAGGCTCTGAGAGAGCCATCTGCGCAAAAGAGTCCCTCAT 712
DB 8441 TCTAGTCAAGTCAAGCTGCGCTGCGCAAGATTCATTCATTAATTAAGTCCCTATTA 8500
QY 713 ATCAACCAAGAGACTCAAAAGACTTGTGCGGAGCAATCCCAATATGACAGACTCTAC 772
DB 8501 ATCAACCAAGAGACTCAAAAGACTTGTGCGGAGCAATCCCAATATGACAGACTCTTA 8560
QY 773 TCTTAGTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832
DB 8561 TCCCTTGTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8620
QY 833 AGACTGCAAGAGCCCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 892
DB 8621 AGACTGCAAGAGCCCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 8680
QY 893 GACCAAGCAACCTTCAAGCAAAATGAGCCAGCTGATGATGATGATGATGATGATGAT 952
DB 8681 AAACGACAAAGTGAACCAAAATGAGCCAGCTGATGATGATGATGATGATGATGATGAT 8740
QY 953 TTGACACTATTTATGACCGCTGAGAGCAAGCAAAATTTGATCAAGCTCCCTCTC 1012
DB 8741 CTGACAAACAACTTATGAGAGCTTGAAGAAATGATGATGATGATGATGATGATGATG 8800
QY 1013 TGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1072
DB 8801 TGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8860
QY 1073 ATCCGTGCTGCTTTTAAACTGCAATATTTCCCTGCTGTAAGCAATTTGGAAGAC 1132
DB 8861 ATTAGAGTCAAGAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATG 8920
QY 1133 AAGTCAAGATATCTTTTGAAGAGTGGCAAGTTCAAGAGATTTGACAGAGGAGG 1192
DB 8921 AATATCAGATATCTCTTTAAGAGATTTGGGGCCGACAGAAATGATGATGATGATGAT 8980
QY 1193 CTGGGCTCTCTTCTGATGATTTTCAATTCAAATTCAGAGAGTGGTGAAGTGCATCC 1252
DB 8981 CTGGGCTCTTACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 9040
QY 1253 TTGGGGGAGTAAATGATGAGCAAGTGTCCGAGCTGCTTCCAA 1297
DB 9041 TTGGAGGAGTAAATGATGAGCTTATGCTTGCAGCTCTTCCAA 9085

RESULT 7
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DOMNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-Fls
US-08-232-463-14

Query Match 3.8%; Score 50; DB 1; Length 7218;
Best Local Similarity 4.2%; Pred. No. 2.1e-05;
Matches 14; Conservative 191; Mismatches 131; Indels 0; Gaps 0;

QY 18 AGAAGAGACGATGACATAGCGCTCAAGAGGAGTGAACCTAAGACCTGAA 77
DB 1381 RR 1332
QY 78 TCATGAGTACTCTTGAGACTGACAAATATTTGACAGAGAGCCTTGAAGAGACTAG 137
DB 1321 RR 1262
QY 138 AAAAACTACAGAGAGCCAGAGAGCTGCCCTCGAGAGAGAGAGCCAGAGATGACTC 197
DB 1261 RR 1202
QY 188 GGCCTTCTAGAAAGAGAGCTGAGAGGTCATATCTAGTGGGAAAAATGAAGCTGACT 257
DB 1201 RR 1142
QY 258 CCGCTGACTGCGAGAGAAAAATGATGAGACCTTGAAAGACTCCAGAGACTTCAAGAGG 317
DB 1141 RR 1082

QY 318 CCAAGATGAGCTGAGACTCAAGCTGCGCCCAAGCTG 353
DB 1081 RR 1046

RESULT 8
US-09-668-313A-10
Sequence 10, Application US/09668313A
Patent No. 6503756
GENERAL INFORMATION:
APPLICANT: Brett P. Monia
APPLICANT: Susan M. Freier
APPLICANT: Jacqueline Wyatt
TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRES.
FILE REFERENCE: PUS-0127
CURRENT APPLICATION NUMBER: US/09/668,313A
CURRENT FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 247
SEQ ID NO 10
LENGTH: 2574
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (218) ... (1891)
US-09-668-313A-10

Query Match 3.4%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 54.7%; Pred. No. 0.00086;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 674 TGGAGAGAGCCATCTGCGCAACAAAGTGCCTACTATATCAACAGAGAGCTCAAGA 733
DB 1733 TGGAGAGAGCTTACAGAGAGTGAATCAAGTACTTATCAACAGAGAGAGAGC 1792
QY 734 ACTTGCTGGAGCATTCCCAAAATGACAGAGCTTACAGCTTTAGCTGAGTAAT 793
DB 1793 ACCTCCTGATCCACCCCGTGATGAGCGGCTGMACTCTCTGTGCAAGAGAGAGTGA 1852
QY 794 GTCAGATTCTCAGCTTATAGAGCTGCCATGAATCCGAAG 834
DB 1853 GAGAGCTGCCAGAGAGCTAACAGACCCGAAAGCTGATG 1893

RESULT 9
US-09-620-312D-69
Sequence 69, Application US/09620312D
Patent No. 6568662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Weinman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunding
APPLICANT: Wang, Dunru
APPLICANT: Wang, Zhimei
APPLICANT: John Tillinghaast
APPLICANT: Dimaac, Radoje T.
TITLE OF INVENTION: No. 656862el Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIORITY APPLICATION NUMBER: 09/552,317

```

? PRIOR FILING DATE: 2000-04-25
? PRIOR APPLICATION NUMBER: 09/488,725
? PRIOR FILING DATE: 2000-01-21
? NUMBER OF SEQ ID NOS: 1105
? SOFTWARE: pc_fl_genes Version 1.0
? SEQ ID NO 69
? LENGTH: 1690
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
?
? FEATURE:
?
? NAME/KEY: CDS
?
? LOCATION: (128)..(1522)
US-09-620-312D-65

```

Query Match	3.3%	Score 42.8;	DB 4;	Length 1690;
Best Local Similarity	47.9%;	Pred. No. 0.0019;		
Matches 156;	Conservative	0;	Mismatches 167;	Indels 3;
				Gaps 1;

Qy	301	CCAGGAACTTCAAGAGGCCACGGATGACCTGACCTCAAGCTGGCCAGCTGAGTGAT	360
Db	55	CAAGAGTTGCACACAGGTGGCGCACGACCTGACGAGAGCTGGCATGGGTTCCAGGAGG	114
Qy	361	CAAGGGA---TCTTGGCAGCCCGTGGGGGAGTCTCTCTATGACTCTCTCCAAATYACT	417
Db	115	GCTGCCACTGGCCATGCACAGACGAGCCAGGGCAAGGTTTGGAGCGGCTCCAGACGACT	174
Qy	418	CGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCCCTCTGAAAGAGACGTGACCACGT	477
Db	175	CAAAAAGAACAGGGGCGTGGCGGGAGATCCAGGCCCATGGGCGCGCCTGGAGGAGGT	234
Qy	478	CAATGACCTTGTCTCGCAGACTTACCACTTTGGGCAATTGACTCTCAACCGTAATPACTAG	537
Db	235	GCTGGAGCGCGGGGGCGGCGGTGGCGTCCCTCGACGCCCCGAGGCGAGGCACTGGCGCG	294
Qy	538	CACTTGGAGAGACTTGAAACCCAGATGAGAACTTCTCAGATGGCCGCTCGAGGACCCGACT	597
Db	295	GGGCGCTGGAGCAGACTTCAGAGCGCCTGGGGCCGAGCTGGGGAGGCTTCCGAGCGCAG	354
Qy	598	CAGGCAGCTGCATGAAGCCCAAGGG	623
Db	355	GCAAGTGTGAGACGCCGCTTCCAGG	380

```

RESULT 10
US-09-368-590-1
/ Sequence 1, Application US/0936850
/ Patent No. 6187563
/ GENERAL INFORMATION:
/ APPLICANT: Solimena, Michele
/ TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
/ TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
/ FILE REFERENCE: 101918-200 (OCR-941)
/ CURRENT APPLICATION NUMBER: US/09/368-590
/ CURRENT FILING DATE: 1999-08-04
/ EARLIER APPLICATION NUMBER: 60/095,657
/ EARLIER FILING DATE: 1998-08-07
/ NUMBER OF SEQ. ID NOS.: 8
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 7812
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(6879)
/ NAME/KEY: unsure
/ LOCATION: (100)...(102)
/ NAME/KEY: unsure
/ LOCATION: (1021)...(1023)
/ NAME/KEY: unsure
/ LOCATION: (2266)...(2268)
/ US-09-368-590-1

```

Query Match	3.3%;	Score 42.8;	DB 3;	Length 7812;
Best Local Similarity	47.9%;	Pred. No. 0.0049;		
Matches 156;	Conservative 0;	Mismatches 167;	Indels 3;	Gaps 1;

QY	300	CCAGGAACCTTCAAGAGCCACAGATAGCTGAGACCTCAAGCTGCGCAAGCTGAAGTAT	360
Db	3639	CAAGAGATTGCACCAAGGTGGCGCACACCTCGACGACGAGCTGGCATGGGTTCAAGACG	3698
QY	361	CAAGGA---TCTCGACAGCCCTGTGGGCATCTCTCATTTGACTCTTCCAAGATCACT	417
Db	3699	GCTGCCACTGGCCATGCAAGACGAGCCAGGCAACGGTTTGCAGGCGGTCCAGCAACAT	3758
QY	418	CGAGAAGCTCAAGCACTTCGAGAGAAATTTGGCCCTCGAAGAGAAAGCTGAGCAGCT	477
Db	3759	CAAAAAGAACCCAGGGCGCTGGCGGCGGAGATCCAGGCGCATGGGCGCGCTGAGAGAGT	3818
QY	478	CAATACCTTGTCTGGCCACGTTACATTTTGGGCATTTGACCTCAACGTTAAACTGAG	537
Db	3819	GCTGAGAGCGCGGGCGCGCTGGGTGCTGGCAAGCCGGAAGCAGAGCAGTGGCGCG	3878
QY	538	CACTCTGGAGACCTGAAACAACAGATGAAGCTTTGACAGTGGCCGCTGAGGACCGAGT	597
Db	3879	GGGCGCTGGAGCAGCTGCAGACCGGCTGGGCGGACCTGGCGGAGGCTGCGAAGCGACCA	3938
QY	598	CAGGACGCTGATGAAAGCCACAGGG	623
Db	3939	GCAAGTGTGAACGCGCTTTCCAGG	3964

RESULT 11
US-08-425-069-3
Sequence 3, Application US/08425069
Patent No. 5728610
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Himman, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5728610th Washington Street
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22046
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425,069
FILING DATE: 19-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-106P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1995 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO

IMMEDIATE SOURCE:
CLONE: p6B
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1785
US-08-425-069-3

Query Match 3.0%; Score 38.6; DB 1; Length 1995;
Best Local Similarity 44.0%; Pred. No. 0.049;
Matches 164; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 11 GCAGTTGAGAGCAGACGATGTACATAGGGCTTCAAGAGGAATTGAACTAAAGAA 70
DB 731 GCAGTGACAGCTCAGAGCCGAGCAGAGACTGACAAACAGGACCCGAGATATGAC 790
QY 71 CCTGTATCATGTAGTACTTTAGACTGTACGAATTTTCTGACAGACAGCTTTGAA 130
DB 791 CAGGACACAGAGACCATGTGACCCGGATGTGCGCTGACAGACACCCCGCAGCAG 850
QY 131 GGAAGTACAGAACTTACCCAGAGCCAGAGAGCTGCTCTGAGAGAGAGCCAGAT 190
DB 851 GACCTGAGAGATATGGCCCTTGACAAACAGACCCGAGAGATATGACACAGACAG 910
QY 191 GTCACTGGCTTTACGAAAGAGCGCTGAGAGTCAATCTGAGTGGAAAAATTGAC 250
DB 911 GACCACTGAGACAGAGAGTGCAGAGCAGCAGCCGAGAGAGACTTGAACAGAGAT 970
QY 251 CTGCACTCCGCTGACTGCGAGAGAAAAATAGATGAGACCTTGAAGACTCCAGAGACT 310
DB 971 TAGGAGTTATGAGACCAAGACCAAGGTCCAGAGAGATATGACCAAGACCAAGATC 1030
QY 311 CAGAGGCCACGAGATGAGCTGAGCTCAAGCTGCGCAGAGCTGAGGTATCAGAGATCC 370
DB 1031 CAGAGAGATATGAGACAGGTAGTGCATCTGACAGCAGACCGCAGACAGACTTGAACAC 1090
QY 371 TGGCAGCCCGTGG 383
DB 1091 AAGGACAGAGAGG 1103

RESULT 12

US-08-317-844B-3
Sequence 3, Application US/08317844B
Patent No. 5989894

GENERAL INFORMATION:

APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming

APPLICANT: Hanman, Michael B.

TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF

NUMBER OF SEQUENCES: 62

CORRESPONDENCE ADDRESSES:

ADDRESSES: Birch, Stewart, Kolasch & Birch

STREET: 301 No. 5989894th Washington Street

CITY: Falls Church

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22046

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/317,844B

FILING DATE: 04-OCT-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 1447-105P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1995 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
IMMEDIATE SOURCE:
CLONE: p6B
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1785
US-08-317-844B-3

Query Match 3.0%; Score 38.6; DB 2; Length 1995;
Best Local Similarity 44.0%; Pred. No. 0.049;
Matches 164; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 11 GCAGTTGAGAGCAGACGATGTACATAGGGCTTCAAGAGGAATTGAACTAAAGAA 70
DB 731 GCAGTGACAGCTCAGAGCCGAGCAGAGACTGACAAACAGGACCCGAGATATGAC 790
QY 71 CCTGTATCATGTAGTACTTTAGACTGTACGAATTTTCTGACAGACAGCTTTGAA 130
DB 791 CAGGACACAGAGACCATGTGACCCGGATGTGCGCTGACAGACACCCCGCAGCAG 850
QY 131 GGAAGTACAGAACTTACCCAGAGCCAGAGAGCTGCTCTGAGAGAGAGCCAGAT 190
DB 851 GACCTGAGAGATATGGCCCTTGACAAACAGACCCGAGAGATATGACACAGACAG 910
QY 191 GTCACTGGCTTTACGAAAGAGCGCTGAGAGTCAATCTGAGTGGAAAAATTGAC 250
DB 911 GACCACTGAGACAGAGAGTGCAGAGCAGCAGCCGAGAGAGACTTGAACAGAGAT 970
QY 251 CTGCACTCCGCTGACTGCGAGAGAAAAATAGATGAGACCTTGAAGACTCCAGAGACT 310
DB 971 TAGGAGTTATGAGACCAAGACCAAGGTCCAGAGAGATATGACCAAGACCAAGATC 1030
QY 311 CAGAGGCCACGAGATGAGCTGAGCTCAAGCTGCGCAGAGCTGAGGTATCAGAGATCC 370
DB 1031 CAGAGAGATATGAGACAGGTAGTGCATCTGACAGCAGACCGCAGACAGACTTGAACAC 1090
QY 371 TGGCAGCCCGTGG 383
DB 1091 AAGGACAGAGAGG 1103

RESULT 13

US-09-220-132-24
Sequence 24, Application US/09220132
Patent No. 6506607

GENERAL INFORMATION:

APPLICANT: Sayjan, Andrew W.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CT

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 24

LENGTH: 7672

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-24

Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpax11.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 48 Row: 6 Column: 7
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: retained intron.

FEATURES

Source

1. 3870
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5274415"
 /tissue_type="Brain, hippocampus"
 /clone_id="NIH_MGC_95"
 /lab_host="DH10B"
 /note="Vector: Bluescript"

BASE COUNT 1321 a 577 c 763 g 1109 t
 ORIGIN

Query Match 28.9%; Score 693.8; DB 11; Length 3870;
 Best Local Similarity 92.7%; Pred. No. 3.5e-135;
 Matches 728; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

1 GGGAGTTGATGAGAGTGAAGTAACTGAGCCGTATCAACAGCTTGAAGA 60
 1143 GGGAGTTGATGAGAGTGAAGTAACTGAGCCGTATCAACAGCTTGAAGA 1202
 61 GATATTCGTCGCTCTTTCTGCTGAGAGACATTGACACAGAGAGATTGTAAT 120
 1203 GATATTCGTCGCTCTTTCTGCTGAGAGACATTGACACAGAGAGATTGTAAT 1262
 121 GATGAGAGTGGTGAAGACCAATTGATTCATGAGGGGTACATGATGATTGAACA 180
 1263 GATGAGAGTGGTGAAGACCAATTGATTCATGAGGGGTACATGATGATTGAACA 1322
 181 GCCCATGAGGGCGGTTGTAATATTCTAACAATTGGAGAGTAAGCTGATTGACAGCA 240
 1323 GCCCATGAGGGCGGTTGTAATATTCTAACAATTGGAGAGTAAGCTGATTGACAGCA 1382
 241 AATTATCAAGAGTGAAGAACTGAACTCAAGAGAGTGAATCTCTTAATTGAAGA 300
 1383 AATTATCAAGAGTGAAGAACTGAACTCAAGAGAGTGAATCTCTTAATTGAAGA 1442
 301 TGGGAATGCTCAGGGTAGCTAGCATGGAACCAAGCAATTGACATGAGTTTAATG 360
 1443 TGGGAATGCTCAGGGTAGCTAGCATGGAACCAAGCAATTGACATGAGTTTAATG 1502
 361 GATTCAGAGATCAAGAACTGAAAGATTGAATGATCGCTAACCAAAAACAGAAAGA 420
 1503 GATTCAGAGATCAAGAACTGAAAGATTGAATGATCGCTAACCAAAAACAGAAAGA 1562
 421 ACAAGAAAATGAGAGAGAGCCCTTGACCTGATCTTGAAGACCTTAACCCCAAGTA 480
 1563 ACAAGAAAATGAGAGAGAGCCCTTGACCTGATCTTGAAGACCTTAACCCCAAGTA 1622
 481 CAACACATTAAGGTGCTTCAAGAAAGATCTGAAACAGACAAAGTCAAGGTCATTTCTC 540
 1623 CAACACATTAAGGTGCTTCAAGAAAGATCTGAAACAGACAAAGTCAAGGTCATTTCTC 1682
 541 ACTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 1683 ACTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1742
 601 GAACACATTAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 1743 GAACACATTAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1802

QY 661 TGGGTTCTTTTACAGACACTATGATTAATCTGACACAGTTCCCTCGACCTGGAAG 720
 Db 1803 TGGGTTCTTTTACAGACACTATGATTAATCTGACACAGTTCCCTCGACCTGGAAG 1862
 QY 721 TTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Db 1863 TTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1922
 QY 781 AAGGA 785
 Db 1923 AAGGA 1927

RESULT 2
 BI729851 834 bp mRNA linear EST 20-SEP-2001
 LOCUS 60334951.1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5',
 DEFINITION mRNA sequence.
 ACCESSION BI729851 GI:15706864
 VERSION BI729851.1 GI:15706864
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 834)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgsbds-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LRAM1908 row: e column: 03
 High quality sequence stop: 796.

FEATURES

Source

1. 834
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:5357162"
 /tissue_type="retina"
 /issue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_94"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dn primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 275 a 173 c 199 g 187 t
 ORIGIN

Query Match 25.2%; Score 605.4; DB 12; Length 834;
 Best Local Similarity 88.6%; Pred. No. 1.1e-116;
 Matches 690; Conservative 0; Mismatches 86; Indels 3; Gaps 3;

QY 628 TGGGCAACATCTGTATGATGACAGAAAGCCGCTGCTTCTTTACAGACACTGATGA 687
 Db 31 TGGGCAACATCTGTATGATGACAGAAAGCCGCTGCTTCTTTACAGAAACTCATGA 90
 QY 688 TTACTGCAACAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 747
 Db 91 TTACTGCAACAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 150
 QY 748 ACACTGCAACATCTGTATGATGACAGAAAGCCGCTGATGAGAAAGCTCTTGAAGACTCCAG 807
 Db 151 ACACTGCAACATCTGTATGATGACAGAAAGCCGCTGATGAGAAAGCTCTTGAAGACTCC-AG 209

QY	608	GGAGTAAAGAGCTGATGAAACATGCGAAGCTCCCAAGTGAATTTGAGCTACACA	867
Db	210	GGAGTCAGAGAGCTGATGAAACATGCGAAGCTCCCAAGTGAATTTGAGCTACACA	269
QY	868	GATGTTTATCAACCTGATGAAACATGCGAAGCTCCCAAGTGAATTTGAGCTACACA	927
Db	270	GATGTTTATCAACCTGATGAAACATGCGAAGCTCCCAAGTGAATTTGAGCTACACA	329
QY	928	GATGATCACTGCTGTTACAAAGAGCTTTGATTAACATGAACTTCAAGTGAAGTGA	987
Db	330	GATGATCACTGCTGTTACAAAGAGCTTTGATTAACATGAACTTCAAGTGAAGTGA	389
QY	988	CGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCAGTTTCAGACAGTGAAGCT	1047
Db	390	CGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCAGTTTCAGACAGTGAAGCT	449
QY	1048	CTGCACCTTTCTCTGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1107
Db	450	TTGCATCTTTCTCTGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	509
QY	1108	CGGACGAGCCTTATGAGAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1167
Db	510	CGTCAAGGACCCATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	569
QY	1168	GCCTTCAAGAGGAAATTTGAAACTTAAGAGCTTGAATCATGAGTACTTGTGAAGCT	1227
Db	570	GCCTTCAAGAGGAAATTTGAAACTTAAGAGCTTGAATCATGAGTACTTGTGAAGCT	629
QY	1228	CGAATATTTTCTGACAGAGAGCTTTGAAAGAGCTAGAGAAATCTC-TACAGAGAGCC	1286
Db	630	AGAAATATTTCTGACAGAGAGCTTTGAAAGAGCTAGAGAAATCTTTCAGAGAGCC	689
QY	1287	AGAGTGCCTCTCTGAGAGAGAGCCAGATGTCTCTGAGAGAGAGAGAGAGAGCT	1346
Db	690	AGAACTGCTCTCTGAGAGAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGCT	749
QY	1347	GGAGTCAATACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1404
Db	750	AGAGTCAATACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	808
RESULT 3			
AK044536			
LOCUS			
DEFINITION	AK044536 3056 bp mRNA linear HTC 05-DEC-2002		
KEYWORDS	Mus musculus adult retina cDNA, RIKEN full-length enriched library,		
SOURCE	close:4930019F21 product:dystrophin, muscular dystrophy, full		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Carninci, P. and Hayashizaki, Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
TITLE	Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to		
MEDLINE	prepare full-length cDNA libraries for rapid discovery of new genes		
PUBMED	Genome Res. 10 (10), 1617-1630 (2000)		
JOURNAL	20499374		
MEDLINE	11042159		
PUBMED			
REFERENCE			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,		
TITLE	Kono, H., Akiyama, J., Nishi, K., Katsunaka, T., Tashiro, H., Itoh, M.,		
JOURNAL	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
MEDLINE	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Aikawa, T., Boro, H., Carninci, P., Fukuda, S., Furuno, W., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hirazawa, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, H., Kondo, F., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ono, N., Ozaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sato, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE

JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/
URL: http://lncm.gsc.riken.go.jp/.

FEATURES

Source

Location/Qualifiers

1..3753

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM.DB:C130016K19"

/db_xref="taxon:10090"

/clone="C130016K19"

/tissue_type="head"

/clone_id="RIKEN full-length enriched mouse cDNA library"

/dev_stage="16 days embryo"

181..3054

/note="unlabeled protein product; dystrophin related protein 2 (WGD MGI:107432, GB|U43520, evidence: EIASIN, 98%, match=807)

putative"

/codon_start=1

/protein_id="BAC38217.1"

/db_xref="GI:26349155"

/translation="MQLVWQGCPTLPRCHEVHAADRFHSSLNTPQPOVRAAV
TTPAPWDGADPCLSPKLVGATGPLEPAMNLCWEIKKSHNLRARLEASD
LSGLQLPRLIIDLWQKDELSQLDGLVALVQKETAAMEVSKGPYS
SVLESAQAFSLQHFELLESSEKSTSPRORIOMSRVWQATVAELMEKLTAR
CVDQRIHETLLELLEIQAMELSTLQAGVATVEPIDLIDLPHEIIOIX
LKEEPSPVADGVKLVNDLHOLAISDVHSMENSRALDINRMKOLQSVAREIKO
LODARHPRGSGOHLSTSYOVMEPFRFNKPYITNHOACTCDHRTMELYTL
ADLNKIFSAIRTKMLRVKALRDLVLTTLALIFENHDAQSEHWDVVEVHC
LTLALYERLEBERGLVNVPLCVDMNLNLNVDGSRGSRMLSPKTIACLCGEV
KEXLOVLEFQVANSQCDQRLHGLALHLEAIQVROLGEVAAFGSNVPSYRCSFR
STGKPIVLEASQFLWNLBQDSQWMLAVLRVITIAOVKTCSCICROCPKGFYR
SLKQFNVDICQCTFLGRASKNKLHPIMEVYITPTTSENMDPITLTKNPSRQY
FSKSPRGYLPVOSVLESDCSPTASSPMIPADTSTRLEHPSRAEENKSPEN
DSLSPDSIDEDQILRHSPITDRERFQQAFCMALESKEKELTIAHLEDEKRI
LQELRLKQHEEAAPLVLEGSAAPTDHRNEBLAEARILRHKSRLETRMQL
EDHNKQLESQLEHRELLOPESDNGSGLASSPSRSGSHPRKGGOTTPTDTE
VADVVGSKQDVSLCLEDIMEKLRHAFPSVSSDVANTLLAS"

3731..3736

/note="putative"

3753

/note="putative"

3753

polya_site

polya_site

3753

/note="putative"

3753

BASE COUNT

927 a 1021 c 938 g 867 t

ORIGIN

Query Match

Best Local Similarity

Matches 888; Conservative 0; Mismatches 551; Indels 0; Gaps 0;

QY	963	CATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCAACATTAGTCCATTGGA	1022
DB	414	CATGAATCTGTGTGGATGAAATATAAAAAAGAGCTCAACACCTCGCGCTGGA	473
QY	1023	ACCAGCTTGTGACCCAGTGAAGCGTCTGCACTTTCTCTGACGAACATCTGTGTGCT	1082
DB	474	GACCTTCTCAGACCTCAAGTGAAGAACTTCAGTCCGCTCCGAGAGTATTTAGCTGCT	533
QY	1083	ACAGCTGAAGATGATGATTAATTAAGCCGAGGACCTTAATGAGGCGACTTCCAGCAGT	1142
DB	534	CAGCCAAAAGATGAGGAGTGTCACTCACTGCTCCCTTGAAGGGAGTGTGCGCTGT	593
QY	1143	TGAGAGCAAGAACGATGATCATAGAGGCTTCAAGAGGAAATTAAGAACCTGT	1202
DB	594	ACAAAGAGGAAGAGAGACATGACAGCTTCAATGAAAGATCAAGTCAAGGGCCCTTA	653
QY	1203	ATTCATGAGTACTCTTGAAGTGTGCAATTAATTCGACAGAGCGCTTGGAGAGACT	1262
DB	654	CATCTCTCTGTGTGATTAATTCGAGGCTTTCTGTCCAGCATTCATTTGAAGAAAT	713
QY	1263	AGAGAACTCTACCAAGAGCCAGAGAGCTGCTCTGAGAGAGAGCCAGAAATGTAC	1322
DB	714	AGAGAGTCTCAATCTGAGAGTAAAGATACCTCCCGACAGAGCGGATTCAAAACCTTAG	773
QY	1323	TGGCTTTCTACAAAGCGGCTGAGAGAGTCAATCTAGTGGGAAAAATTGAACCTGCA	1382
DB	774	TGCGTTTATGAGAGAGGAGCAAGGTGCGCAAGTGAAGTGTGGAGAACTGACAGCCG	833
QY	1383	CTCCGCTAGTGGCAGAGAAAAATGATGAGACCTTGAAGACCTCCAGAACTTCAAGA	1442
DB	834	CTGTGTATATAGATATCCCATTTGAGACATCTGTGAAGATCTATTGAGATCCAGG	893
QY	1443	GCCCAAGATGAGTGTGACCTCAAGCTGCGCAAGCTGAGTGAATCAAGGATCTGCGCA	1502
DB	894	GCCATGAGAGGATTTGACAGTACTTTGACCCAAAGAGAGGAGGAGCCAGCGTGGGA	953
QY	1503	GCCCGTGGGCAATCCCTCATTTGACTCTCTCCAGAGTACCTCGAAGAGCCAGGCACT	1562
DB	954	GCCATGAGAGATCTCTTATTTGATCCCTCCAGAGATATCCAGGCATCAAGCTATT	1013
QY	1563	TCGAGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACCTCAATGACTTGTGCGCA	1622
DB	1014	CAAAGAAATTCCTCTCTGAAAGAGAGGAGTGAATGATGATGATGATGATGATGATGAT	1073
QY	1623	GCTTACCATTTGGGCAATTCAGCTCTCAACCTATTAACCTCAGCACTCTGGAAGACTGAA	1682
DB	1074	GCTTGCATCTCTGATGAGCACTGTCAATGAGAAATTCAGGGCTCTGGAACAGATCAA	1133
QY	1683	CACCAAGATGAGAGCTTGTGAGAGTGGCGGTGAGAGACCGATGAGCGAGCTGATGAAGC	1742
DB	1134	CATCGGTGGAAGACGCTCCAGGTGTCAAGTGTGAGAGGCTTAAGCAATCCAGAGATGC	1193
QY	1743	CCACAGGACTTTGGTCCAGCATTCAGCACTTTTCCAGCTGTGTCCAGGGTCCCTG	1802
DB	1194	CCACCGGAGACTTTGGGCTGGGTGACAGCACTTCTCCCTCTGCGCAAGTTCCTTG	1253
QY	1803	GGAAGAGCCATTCGCAAGAGGCGCTTATATATCAACAGAGAGACTCAAAAC	1862
DB	1254	GGAAGAGCAATTTCTCCAAATTAAGTCCCTTACTACATCAACACAGGCTCGAGACAC	1313
QY	1863	TTGCTGGACCATCCCAATATGACAGAGCTTACAGTCTTTAGCTGAAATTAATGT	1922
DB	1314	ATGCTGGACCATCTTAAGATGATGATGATTAACAAACCTTAAGCCGATGGAACAAT	1373
QY	1923	CAGATTCAGCTTATAGAGATTCGCAATGAACCTCCGAAGACTGAGAGAGGCGCTTGT	1982
DB	1374	TAACTTCAGCTTACGCACTGAGAGCTCCGAGATCCGAGATCCGAGAGGCGCTTGT	1433
QY	1983	GGATCTTGAAGCTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT	2042
DB	1434	GGATGTGATCTTAACTTAAGCTGTGAGATGATGATGATGATGATGATGATGATGAT	1493

QY 2043 TGACCGCCCATGATATCTGCGATTTAATTGTTTGAACCACTATTATGACCGCT 2102
DB 1494 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553
QY 2103 GGACCAAG 2162
DB 1554 GGAG 1613
QY 2163 GCTGCTGAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2222
DB 1514 GCTCCTGAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673
QY 2223 TGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2282
DB 1674 TGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1733
QY 2283 AGTGGCAAGTTCACAGAGATTTTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 2342
DB 1734 AGTGGCAAGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1793
QY 2343 TATCCAAATTCAG 2401
DB 1794 CATTCAGTGGCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1852

RESULT 5

AL556247

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL556247 824 bp mRNA linear EST 31-MAY-2003
AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
AL556247 CDNA clone CS0D001YB17 5-PRIME, mRNA sequence.
AL556247.2 GI:31278051
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 824)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12898746.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/InvitrogenCorporation 1600
Faraday Avenue Genoscope sequence ID: CS0D001CA09QPL.
Location/Qualifiers
1..824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D001YB17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-cligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 211 a 210 c 206 g 197 t.
ORIGIN

Query Match 23.2%; Score 557; DB 9; Length 824;
Best Local Similarity 100.0%; Pred. No. 1.6e-106;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1845 CCACGAGACTCAACACTGCTGAGACATCCCAAAATAGACAGAGCTCTACAGTCTT 1904

DB 195 CCACGAGACTCAACACTGCTGAGACATCCCAAAATAGACAGAGCTCTACAGTCTT 254
QY 1905 AGCTGACCTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1964
DB 255 AGCTGACCTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 314
QY 1965 GAG 2024
DB 315 GAG 374
QY 2025 GCACAG 2084
DB 375 GCACAG 434
QY 2085 CACTATTTATGACCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2144
DB 435 CACTATTTATGACCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
QY 2145 GAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2204
DB 495 GAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 554
QY 2205 TGTCTGCTTTTAAAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2264
DB 555 TGTCTGCTTTTAAAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 614
QY 2265 CAGATACCTTTTCAAG 2324
DB 615 CAGATACCTTTTCAAG 674
QY 2325 CTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2384
DB 675 CTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 734
QY 2385 GGGCAGTAACTTGAAGC 2401
DB 735 GGGCAGTAACTTGAAGC 751

RESULT 6
CB991394 801 bp mRNA linear EST 01-MAY-2003
LOCUS
DEFINITION
AGENCOURT 13627932 NIH-MGC 148 Homo sapiens cDNA clone
IMAGE:30336570 5', mRNA sequence.
CB991394
CB991394.1 GI:30285818
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 801)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (NLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/NIH at:
http://image.llnl.gov
plate: NDAM360 row: d column: 19
High quality sequence stop: 621.
Location/Qualifiers
1..801
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source
1..801
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

		/tissue_type="pre-ecIamptric placenta" /lab_host="DH10B Tona" /clone_idb="Nih_MGC_148" /note="Organ: Placenta; Vector: pBluescriptR, Site: 1: all-XhoI; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTNNNNNN-3', size-selected for average insert size 2.3 kb and normalized to ROL 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NTM/NHGRI, National Institutes of Health). Note: this is a Nih_MGC Library."			
BASE COUNT	204 a	210 c	199 g	188 t	
ORIGIN					
Query Match	23.1%	Score 555.4;	DB 14;	Length 801;	
Best Local Similarity	99.8%;	Pred. No. 3.4e-106;			
Matches 556;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1845	CCAGGAGCTCAAAACACTGTGTGGAGCACATCCAAATTGAAGAGCGCTCAACAAGCTTT	190		
Dd	140	CCAGGAGCTCAAAACACTGTGTGGAGCACATCCAAATTGAAGAGCGCTCAACAAGCTTT	199		
QY	1905	AGCTGACCTGATAATATGTTCAGATTCTCAGCTTAATAGAAGCTGCCAATGAAGCTCCGAAGACT	1964		
Dd	200	AGCTGACCTGATAATATGTTCAGATTCTCAGCTTAATAGAAGCTGCCAATGAAGCTCCGAAGACT	259		
QY	1965	GCAGAGAGCCCCCTTGCTTGATCTCTTGAGCCTGTCACTGATGTGATGCTTGAGACCA	202		
Dd	260	GCAGAGAGCCCCCTTGCTTGATCTCTTGAGCCTGTCACTGATGTGATGCTTGAGACCA	319		
QY	2025	GCACAACTCCAGCAAATATGACACACCACATGATATCTGACAGATTATTAATGTTTGAC	2084		
Dd	320	GCACAACTCCAGCAAATATGACACACCACATGATATCTGACAGATTATTAATGTTTGAC	379		
QY	2085	CACATTTTATGACCGCCTGAGAACAGACACAAATTGGTCAACGTCCTCTCTGCGT	2144		
Dd	380	CACATTTTATGACCGCCTGAGAACAGACACAAATTGGTCAACGTCCTCTCTGCGT	439		
QY	2145	GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATAGGGAGAGCAACAGGAGAGATCG	2204		
Dd	440	GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATAGGGAGAGCAACAGGAGAGATCG	499		
QY	2205	TGTCCTGTCTTTAAAACCTGGCATCTTCCCTGTGTAAAGACATTGTGAAACAAGTA	2264		
Dd	500	TGTCCTGTCTTTAAAACCTGGCATCTTCCCTGTGTAAAGACATTGTGAAACAAGTA	559		
QY	2265	CAGATACCTTTTCAAGCAAGTGGGCAAGTTTCAACAGAGATTTTGTGACCGCAGGCTGCGG	2324		
Dd	560	CAGATACCTTTTCAAGCAAGTGGGCAAGTTTCAACAGAGATTTTGTGACCGCAGGCTGCGG	619		
QY	2325	CTCTCTTCTGATGATGTTCTATCCAATTCCAAAGACAGTTGGGTGAAGTTGATCCTTTGG	2384		
Dd	620	CTCTCTTCTGATGATGTTCTATCCAATTCCAAAGACAGTTGGGTGAAGTTGATCCTTTGG	679		
QY	2385	GGGCACTAACATTGAGC 2401			
Dd	680	GGGCACTAACATTGAGC 696			
RESULT 7					
LOCUS	BX365572	1098 bp	mRNA	linear	EST 05-MAY-2003
DEFINITION	BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens				
ACCESSION	CDDA Clone CS0DK001YB21 5-PRIME, mRNA sequence.				
VERSION	BX365572 . GI:30366927				
KEYWORDS	EST .				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1098)	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	La M.B., Gruber C., Jessee J. and Polyes, D.	Full-length cDNA libraries and normalization	Unpublished
1		Genoscope		Contact: Genoscope
1		Genoscope - Centre National de Sequencage		Genoscope - Centre National de Sequencage
1		BP 191 91006 Evry cedex - France		BP 191 91006 Evry cedex - France
1		Email: seq@genoscope.cns.fr		Web : www.genoscope.cns.fr
1		Library was constructed by life Technologies, a division of		Invitrogen. This sequence belongs to sequence cluster 1955.r
1		Contact : Feng Liang Email : fliang@lifetech.com URL :		http://fulllength.invitrogen.com/Invitrogen Corporation 1600
1		Faraday Avenue Genoscope sequence ID : CS1AK001ZB11Q1.		Location/Qualifiers
1		1..1098		
1		/organism="Homo sapiens"		
1		/mol_type="mRNA"		
1		/db_xref="taxon:9606"		
1		/clone="CS0DK001YB21"		
1		/cell_type="HELA CELLS COT 25-NORMALIZED"		
1		/cell_line="HELA"		
1		/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"		
1		/note="1st strand cDNA was primed with a NotI-oligo (dT)		
1		primer. Five prime end enriched, double-strand cDNA was		
1		digested with Not I and cloned into the Not I and EcoR V		
1		sites of the pCMVSPORT 6 vector. Library was normalized."		
BASE COUNT	286 a	271 c	268 g	262 t 11 others
ORIGIN				
Query Match	23.1%	Score 555.4;	DB 13;	Length 1098;
Best Local Similarity	99.8%	Pred. No. 3.6e-106;		
Matches 556;	Conservative	0;	Mismatches 1;	Indels 0;
			Gaps 0;	
QY	1845	CCACGACACCTCAACCACTGCTGGAGCACTCCCAATGACAGAGCTTCAAGTCTTT	1904	
DB	204	CCACGACACCTCAACCACTGCTGGAGCACTCCCAATGACAGAGCTTCAAGTCTTT	263	
QY	1905	AGCTGACCTGAATTAATGTCAAGATTCTCAGCTTTAAGAGCTGCCAATCTCCGAAGACT	1964	
DB	264	AGCTGACCTGAATTAATGTCAAGATTCTCAGCTTTAAGAGCTGCCAATCTCCGAAGACT	323	
QY	1965	GCAAGAAAGCCCTTGTCTGATGATCTCTGAGCCCTGACGCTGACATGATGCTTGGACCA	2024	
DB	324	GCAAGAAAGCCCTTGTCTGATGATCTCTGAGCCCTGACGCTGATGATGCTTGGACCA	383	
QY	2025	GCAACAACCTCAAGCAAAATGACAGAGCCCATGATGATCTCAGATTATTAATTGTTGAC	2084	
DB	384	GCAACAACCTCAAGCAAAATGACAGAGCCCATGATGATCTCAGATTATTAATTGTTGAC	443	
QY	2085	CACATTATTATGACCGCTGGAGCAAGGCAACACAAATTGGTCAAGTCCCTCTGTGCT	2144	
DB	444	CACATTATTATGACCGCTGGAGCAAGGCAACAAATTGGTCAAGTCCCTCTGTGCT	503	
QY	2145	GGATATGTGCTGAACCTGCTGCTGAATGTTTATGATACGAGCAAGAGGAGATCCG	2204	
DB	504	GGATATGTGCTGAACCTGCTGCTGAATGTTTATGATACGAGCAAGAGGAGATCCG	563	
QY	2205	TGTCTGTCTTTTAAACTGGACATATTTCCCTGTGTAAGACACATTTGGAAAGCAAGTA	2264	
DB	564	TGTCTGTCTTTTAAACTGGACATATTTCCCTGTGTAAGACACATTTGGAAAGCAAGTA	623	
QY	2265	CAGATACCTTTGAAGAGAGTGGCAAGTTTGAAGCAAGGAGAGGAGG	2324	
DB	624	CAGATACCTTTGAAGAGAGTGGCAAGTTTGAAGCAAGGAGAGGAGG	683	
QY	2325	CTCTCTTCATGATTTCTATCCAAATTCGAAGCAGTTGGAGTGAAGTTCATCTTTGG	2384	
DB	684	CTCTCTTCATGATTTCTATCCAAATTCGAAGCAGTTGGAGTGAAGTTCATCTTTGG	743	
QY	2385	GGGCAATTAACATTGAGC 2401		
DB	744	GGGCAATTAACATTGAGC 760		

RESULT 8
 B0640063
 LOCUS
 DEFINITION he23g04.Y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
 Homo sapiens cDNA clone he23g04 5', mRNA sequence.
 ACCESSION B0640063
 VERSION B0640063.1
 KEYWORDS GI:21764522
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 620)
 Wislow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
 Expressed sequence tag analysis of human retina for the NEIRBank Project: Retbindin, an abundant, novel retinal cDNA and alternative splicing of other retina-preferrd gene transcripts
 Mol. Vis. 8 (4) (2002) In press
 TITLE
 JOURNAL Contact: Wislow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: graham@helix.nih.gov
 Plate: 23 row: 9 column: 04
 Seq primer: M13RP1 reverse primer (ABI).
 FEATURES
 source
 1..620
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="he23g04"
 /tissue_type="Retina"
 /dev_stage="Adult"
 /lab_host="EMD10B"
 /clone_idb="Human Retina cDNA (Un-normalized, unamplified)
): hd/he"
 /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue was dissected from two 80 year old donors with no observed eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the Superscript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTCTAGATCGGCGCGCC(T)15-3'. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."
 BASE COUNT 165 a 148 c 144 g 163 t
 ORIGIN
 Query Match 22.7%; Score 546; DB 13; Length 620;
 Best Local Similarity 100.0%; Pred. No. 3.1e-104;
 Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1856 AAACAATTGCTGGAGACATCCAAAATGACAGAGCTTACCACTTTAGCTGACCTGA 1915
 DB 1 AAACAATTGCTGGAGACATCCAAAATGACAGAGCTTACCACTTTAGCTGACCTGA 60
 QY 1916 AATATGAGATTCCTCAGCTTATAGACTGCATGAACTCCGAGAGCTGAGAGAGCCC 1975
 DB 61 AATATGAGATTCCTCAGCTTATAGACTGCATGAACTCCGAGAGCTGAGAGAGCCC 120
 QY 1976 TTTGCTTGATCTCTTGAGCTGTAGCTGATGATGCTTGGACCAAGCAACTCA 2035
 DB 121 TTTGCTTGATCTCTTGAGCTGTAGCTGATGATGCTTGGACCAAGCAACTCA 180

QY 2036 AGCAATATGACAGCCCATGATATCTCCGAGATTTATTTGTCACACTATTATG 2095
 DB 181 AGCAATATGACAGCCCATGATATCTCCGAGATTTATTTGTCACACTATTATG 240
 QY 2096 ACCGCTGAGAGCAAGAGCAACAATTTGTCACAGCTCCCTCTGCTGATATGTC 2155
 DB 241 ACCGCTGAGAGCAAGAGCAACAATTTGTCACAGCTCCCTCTGCTGATATGTC 300
 QY 2156 TGAATCTGCTGCTGATATTTATGATAGCGGACCAAGAGAGATCCGTCCTCT 2215
 DB 301 TGAATCTGCTGCTGATATTTATGATAGCGGACCAAGAGAGATCCGTCCTCT 360
 QY 2216 TTAAGATGAGCATCTTCCCTGCTGTAAGACATTTGGAAGACAGATGATACCTT 2275
 DB 361 TTAAGATGAGCATCTTCCCTGCTGTAAGACATTTGGAAGACAGATGATACCTT 420
 QY 2276 TGAAGCAAGTGGCAAGTTCAACAGAGATTTGTACACAGAGCTGGGCTCTCTGC 2335
 DB 421 TGAAGCAAGTGGCAAGTTCAACAGAGATTTGTACACAGAGCTGGGCTCTCTGC 480
 QY 2336 ATGATTCATCCAAATTCGACAGACAGTGGGTGAAGTTGCATCTTGGGGACATACA 2395
 DB 481 ATGATTCATCCAAATTCGACAGACAGTGGGTGAAGTTGCATCTTGGGGACATACA 540
 QY 2396 TTGAGC 2401
 DB 541 TTGAGC 546

RESULT 9
 B0719710 770 bp mRNA linear EST 08-MAY-2001
 LOCUS B0719710
 DEFINITION 602690430F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4822807 5',
 mRNA sequence.
 ACCESSION B0719710
 VERSION B0719710.1
 KEYWORDS GI:13998897
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 770)
 NIH-MGC http://mgi.mgi.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shizaki Toshitaki and Piero Carninci (RIKEN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL0731 row: 1 column: 08
 High quality sequence stop: 767.
 FEATURES
 source
 1..770
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4822807"
 /lab_host="DH10B"
 /clone_idb="NIH_MGC_97"
 /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtagag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National

VERSION	AK087829.1 GI:26352777
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamamoto, I., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kado, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Knehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schmitt, L. M., Staahl, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyokawa, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshak-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	6 (bases 1 to 1298)
REFERENCE	6
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P., Furuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, N., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, K., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

JOURNAL	Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-APR-2002) Yoshida Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
FEATURES	Source
FEATURES	Location/Qualifiers
FEATURES	1..1298
FEATURES	/organism="Mus musculus"
FEATURES	/mol_type="mRNA"
FEATURES	/strain="C57BL/6J"
FEATURES	/db_xref="PANTOM:DB:E330026B12"
FEATURES	/db_xref="taxon:10090"
FEATURES	/clone="E330026B12"
FEATURES	/sex="female"
FEATURES	/tissue_type="ovary"
FEATURES	/clone_lib="RIKEN full-length enriched mouse cDNA library"
FEATURES	/dev_stage="2 days pregnant adult"
FEATURES	120..1142
FEATURES	/note="unlabeled protein product; dystrophin, muscular dystrophy (MCD) MGI:94905, GB NM_007868, evidence: BLASTN, 100%, match=999)
FEATURES	putative"
FEATURES	/codon_start=1
FEATURES	/protein_id="BAC40019.1"
FEATURES	/db_xref="GI:26352778"
FEATURES	/translation="MRHKGHEETGTCQNDHPKMTLEYOSLADINVRFSAYRTMKLRLQKACIDILSLSAACDALDQDNKQNDPMDIOLITITVYRLBENHNNLVLPICVDKMLNLNVDTGRGRIRVLSFKIGISLKAHEDKRTLFKQVASTGFCDOIRGLILNDSIQIPROGEVASFGSNIIEPVSQCFQPNKRPTEALFLDMKRLSPSMVWLPVLRVVAALFPAKAKNICEKPIIGFRSLVGFNYDICOSEFSGRVAAGHKHGVYVCEPTTSGEDVRFAKVLKPKFKRYFAHGPVGLVQTVLEGDNNET"
FEATURES	BASE COUNT
FEATURES	343 a 332 c 296 g 327 t
FEATURES	ORIGIN
FEATURES	Query Match
FEATURES	Best Local Similarity 93.9%; Score 502.6; DB 11; Length 1298;
FEATURES	Matches 523; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
FEATURES	1845 CCACGAGCTCAACAACTTCTGAGACCAATCCCAAAATGACAGAGCTTACCGCTCTT
FEATURES	140 CCACGAGCCCAACCACTTGTGGAGACACCCCAAAATGACAGAGCTTACCGCTCTT
FEATURES	1905 AGCTGACCTGAATATGTCAGATTCTCAGCTTATAGACATGCCATGAACTCCGAAACT
FEATURES	200 AGCTGACCTGAATATGTCAGATTCTCAGCTTATAGACATGCCATGAACTCCGAAACT
FEATURES	1965 GCGAAGGCGCTTGTGCTGATCTCTTGAAGCTGTGAGCTGACATGATGCTTGGACCA
FEATURES	260 CCAGAGGCGCTTGTGCTGATCTCTTGAAGCTGTGAGCTGACATGATGCTTGGACCA
FEATURES	2025 GCGAAGCTCAAGCAAAATGACAGCCCATGATATCCTGAGATTATTAATGTTGAC
FEATURES	320 GCGAAGCTCAAGCAAAATGACAGCCCATGATATCCTGAGATTATTAATGTTGAC
FEATURES	2085 CACTATTATGACCGCTGAGCAAGACCAACAATTGGTCAACGCTCTCTGCTGCT
FEATURES	380 TACAATTATGATGCTGAGCAAGACCAACAATTGGTCAACGCTCTCTGCTGCT

[illegible]

Query Match	20.9%	Score 502.6	DB 11	Length 2135
Best Local Similarity	93.9%	Pred. No. 4.8e-95		
Matches 523	Conservative 0	Mismatches 34	Indels 0	Gaps 0
QY	1845	CCACAGAGCTCAAAAGCTGTGGGAGCAACCCAAATGACAGAGCTTACAGTCTT	1964	
DB	144	CCACAGAGCTCAAAAGCTGTGGGAGCAACCCAAATGACAGAGCTTACAGTCTT	203	
QY	1905	AGCTGACCTGATAATATGTCAGATTCCTGAGTTATGAGCTGCCATGAACTCCGAGACT	1964	
DB	204	AGCTGACCTGATAATATGTCAGATTCCTGAGTTATGAGCTGCCATGAACTCCGAGACT	263	
QY	1965	GAGAGAGGCGCTTGTCTTGTGATCTTGTGAGCTGTGACCTGATGATGCTTGGACCA	2024	
DB	264	CCAGAGAGGCGCTTGTCTTGTGATCTTGTGAGCTGTGACCTGATGATGCTTGGACCA	323	
QY	2025	GCACAGCTCAAGCAAAATGACACAGCCAGGATGATCTGAGATTTATTTATTTGTTGAC	2084	
DB	324	GCACAGCTCAAGCAAAATGACACAGCCAGGATGATCTGAGATTTATTTATTTGTTGAC	383	
QY	2085	CACATATTATGACCGCTGAGAGCAAGCAACCAATTGTCACGCTCTCTCTGCT	2144	
DB	384	TACATATTATGACCGCTGAGAGCAAGCAACCAATTGTCACGCTCTCTCTGCT	443	
QY	2145	GGATATGTCCTGAGCTGCTGCTGATTTATGATACGGAGCAACAGGAGAGATCCG	2204	
DB	444	GGATATGTCCTGAGCTGCTGCTGATTTATGATACGGAGCAACAGGAGAGATCCG	503	
QY	2205	TGTCTGCTTTTAAATGAGCAGATTCCTGCTGTATTAACACATTTGTAAGACAATA	2264	
DB	504	TGTCTGCTTTTAAATGAGCAGATTCCTGCTGTATTAACACATTTGTAAGACAATA	563	
QY	2265	CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACAGCGGAGCTGG	2324	
DB	564	CAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACAGCGGAGCTGG	623	
QY	2325	CCTCTCTGATGATTCATGCAATTCCTCAAGACAGTGGGTGAAGTGCATCTTGG	2384	
DB	624	TCTTCTTGTGATTCATGCAATTCCTCAAGACAGTGGGTGAAGTGCATCTTGG	683	
QY	2385	GGGAGTAACATTTGAGC 2401		
DB	684	GGGAGTAACATTTGAGC 700		
RESULT 13				
AK036936				
LOCUS	AK036936	4437 bp	mRNA	linear HTC 05-DEC-2002
DEFINITION	Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930028B14 product:dystrophin, muscular dystrophy, full insert sequence.			
ACCESSION	AK036936			
VERSION	AK036936.1	GI:26085504		
KEYWORDS	HTC; Cap trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			

AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsuna, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiya, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaio, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Nombarelli, P., Nordone, P., Ring, B., Rungwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohzuki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 4437)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Aikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, Y., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct SubMISSION
JOURNAL	Submitted (16-Jul-2001) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsuyumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
FEATURES	Please visit our web site for further details.
source	URL: http://genome.gsc.riken.go.jp/ URL: http://phantom.gsc.riken.go.jp/ Location/Qualifiers 1. 4437

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9930028B14"
/db_xref="taxon:10090"
/clone="9930028B14"
/sex="female"
/tissue_type="vagina"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
29..1968
/misc_feature
/notes="dystrophin, muscular dystrophy (MGD) [MGI:94909,
GB|NM_007868, evidence: BLASTN, 100%, match=999]
putative"
4417..4422
polya_signal
/notes="putative"
4437
polya_site
/notes="putative"
BASE COUNT 1300 a 932 c 885 g 1320 t
ORIGIN

```

```

Query Match      20.9%; Score 502.6; DB 11; Length 4437;
Best Local Similarity 93.9%; Pred. No. 5,3e-95;
Matches 523; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

```

```

QY 1845 CCACGAGACTCAACAACTTGGTGGAGCCATCCAAATGACAGAGCTCTACCACTTT 1904
DB 49 CCACGAGACCCAAACACACTGTGGGACACACCCCAATGACAGCTCTACCACTTT 108
QY 1905 AGCTGACCTGAATATGTATATCTCAGCTTATAGAGCTGCCATGAATCCGAGACT 1964
DB 109 AGCTGACCTGAATATATGTCAGGTTCTCCGCGTATAGAGCTGCCATGAATCCGAGAGCT 168
QY 1965 GCAGAAAGGCCCTTGGTGGATCTCTTGGAGCTGTCAGCTGCATGATGATGCTTGAACCA 2024
DB 169 CCAGAAAGGCCCTTGGTGGATCTCTTGGAGCTGTCAGCTGCATGATGATGCTTGAACCA 228
QY 2025 GCACAACTCAAGCAAAATGACACCCCATGATATCTGCAATATTAATTTGTTGAC 2084
DB 229 GCACAACTCAAGCAAAATGACACCCCATGATATCTGCAATATTAATTTGTTGAC 288
QY 2085 CACTATTTATGACCGCGTGGAGCAAGAGACAAATTTGTCACACGCTCTCTGCT 2144
DB 289 TACAATTTATGATGCTGTGGAGCAAGAGACAAATTTGTCACACGCTCTCTGCT 348
QY 2145 GGATATGTCGTGAATCTGCTGCTGATATGATACGGGACGAAACAGGAGAGATCCG 2204
DB 349 GGATATGTCGTGAATCTGCTGCTGATATGATACGGGACGAAACAGGAGAGATCCG 408
QY 2205 TGTCTGTCTTTAAACCTGGGATCATTTCCCTGTGTAAGCACTTTGGAAGCAAGTA 2264
DB 409 TGTCTGTCTTTAAACCTGGGATCATTTCCCTGTGTAAGCACTTTGGAAGCAAGTA 468
QY 2265 CAGATACCTTTCAAGCAAGTGGCAAGTCAACAGATTTTGTGACCGCAGGAGCTGGG 2324
DB 469 CAGATACCTTTCAAGCAAGTGGCAAGTCAACAGATTTTGTGACCGCAGGAGCTGGG 528
QY 2325 CCTCTTCTGATGATTTTATTCACAAATTTCCAAAGACAGTGGGTAAGTGCATCTTTGG 2384
DB 529 TCTTCTTCTGATGATTTTATTCACAAATTTCCAAAGACAGTGGGTAAGTGCATCTTTGG 588
QY 2385 GGGCAGTAAACATTTGAGC 2401
DB 589 GGGCAGTAAACATTTGAGC 605

```

```

RESULT 14      1384 bp      mRNA      linear      HTC 07-DEC-2002
LOCUS          AK075809
AC075809       Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
DEFINITION     enriched library, clone:1810074B05 product:dystrophin, muscular
                dystrophy, full insert sequence.
ACCESSION      AK075809

```

```

VERSION
AK075809.1 GI:26344655
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
MUS musculus
ORGANISM
REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Komno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Smi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tazaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Atakawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gajobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsumoto, Y., Niki, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombere, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weltz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S.,
and Hayashizaki, Y.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409 (6821), 685-690 (2001)
MEDLINE
21085660
PUBMED
11217851
REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
PUBMED
12117851
REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akahira, S., Akiyama, T., Aono, H., Arai, A.,
Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hirakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komno, H., Kouda, M.,
Koya, S., Kuritara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,

```



```
Db      13 AGAATATCATACTTCTTGSGGAACATTCACTAAACATGAGTAGAGCAAGGCTG 72
Qy      660 CTGGGTTCTTTTACAAAGACACTCATGATTACTGCAACAGTTCCCTGAGACTGGAAAA 719
Db      73 CTGGGAC-----AGAACTCATAGATTACTGACAGCAGTCCCTCTGACCTGAGAA 125
Qy      720 GTTCTTGCTGCG-CTTACAGAAAGTGAAC--AACTGCCATGTCTTACAGAGATGCTAC 776
Db      126 GTTCTCTTCTTGAGATTACGAGAAAGCAAGAGCTGCCAATGTCTTACAGAGAGCTTC 185
Qy      777 CCG-7AAGGAAAGGCTCTAGAAAGCTCCAAAGG--AGTAAAGAGCTGATGAACAATG 833
Db      186 CCGTAAGGAGAGCTCTAGCAAGACTCAGGGGAGTCAAGAGCTGATGAACAATG 245
Qy      834 GCAAGCTCTCCAGAGTGA--TTGAAGCTCACAGATGTTTATCACAACTGAGTGA 892
Db      246 GCAAGTCTCCAGAGAAAGTGAAGTCAACAGATATCTATCACAACTTGTATGA 305
Qy      893 ACAGCCAAAAATCCTGAGATCCCTGAAAGTTCCGATGATGCACTCTGTTACAAAGAC 952
Db      306 ATGGCCAAAAATCCTGAGATCCCTGAAAGTTCCGATGATGCACTCTGTTACAAAGAC 365
Qy      953 GTTGTATATCATGAACCTTCAAGTGAAGTGAACCTTGGAAAAAGTCTCTCAACATTAGGT 1012
Db      366 GTTGTATATCATGAATTCAGAGTGAAGTGAACCTTCAAGAAAAAGTCTCTCAACATTAGGT 425
Qy      1013 CCATTTGAGCCAGTCTGACCAAGTGAAGGCTGACCTTCTCTGAGAGAACTTC 1072
Db      426 CCATTTGAGCAAGTCTGACCAAGTGAAGGCTTGTGATCTTCTCTGAGAACTTC 485
Qy      1073 TGGTGTGCTACAGCTGAAGATGATGAATTAAAGCCGAGGCACTTATGAGGCGACT 1132
Db      486 TTGTTGCTACAGCTGAAGATGATGAACCTGAGCCGTCAGGCACTCATCGTGTGAT 545
Qy      1133 TTCCAGCAGTTCAGAAAGCA-GAACATGTACATAGGCTTCAAGAGGGAATTGAAACT 1191
Db      546 TTCCAGCAGTTCAGAAAGCAAGATGATATACATAGGCTTCAAGAGGGAATTGAAACT 605
Qy      1192 AAAGAACTGTATCATGAGTACTCTGAGACTGTACGATATTTCTGACAGAGAGGCT 1251
Db      606 AAAGAACTGTATCATGAGTACTCTGAGACTGTGAGATATTTCTGACAGAGAGGCT 665
Qy      1252 TTGGAAGGACTGAGAAACTTACCAAGAGCCAGAGAGCTGCTCTGAGAGAGAGCC 1311
Db      666 TTGGAAGGACTGAGAAACTTACCAAGAGGCCAGAGAACTGCTCTGAGAGAGAGCT 725
Qy      1312 CAGAAATGCACTCGGCTTCTAGCA 1335
Db      726 CAGAAATGCACTCGGCTTCTAGCA 749
```

Search completed: February 2, 2004, 06:03:12
Job time : 4507.46 secs

Db 1180 GCCCATCAGGCGCGGTTGTAATATCTACATATGGAGATGAGTATGAGTGAACAGGA 1239
QY 241 AATATATGAGAGATGAGAGAACTGACATACAGAGCAGATGATCTCTAATTCAGGA 300
Db 1240 AATATATGAGAGATGAGAGAACTGAGATACAGAGCAGATGATCTCTAATTCAGGA 1239
QY 301 TGGGATGCTCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 1300 TGGGATGCTCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
QY 361 GATCTCCAGATTCAGAACTGAAAGATTTGATGATGATGATGATGATGATGATGATGATG 420
Db 1360 GATCTCCAGATTCAGAACTGAAAGATTTGATGATGATGATGATGATGATGATGATGATG 1419
QY 421 ACAAGGAAATGAGAGAGAGGCTCTTGAGCTGATCTTGAGAGCTTAAAGCGCAAGTA 480
Db 1420 ACAAGGAAATGAGAGAGAGGCTCTTGAGCTGATCTTGAGAGCTTAAAGCGCAAGTA 1479
QY 481 CAACAACTAAGAGTCTTCAAGAGATCTAGAACAGAACAAAGTCAGGCTCAATCTCTC 540
Db 1480 CAACAACTAAGAGTCTTCAAGAGATCTAGAACAGAACAAAGTCAGGCTCAATCTCTC 1539
QY 541 ACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 1540 ACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599
QY 601 GAACAACTTAAGATTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 1600 GAACAACTTAAGATTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1659
QY 661 TGGCTCTTTTCAAGACATCATGATTTACTGCAACAGTCCCTCGAGCTGAGTGAAG 720
Db 1660 TGGCTCTTTTCAAGACATCATGATTTACTGCAACAGTCCCTCGAGCTGAGTGAAG 1719
QY 721 TTTCTGCTGCTTACAGAGCTGAAACAACCTGCAATGCTTACAGATGCTAACCCGT 780
Db 1720 TTTCTGCTGCTTACAGAGCTGAAACAACCTGCAATGCTTACAGATGCTAACCCGT 1779
QY 781 AAGGAAAGGCTCTTGAAGATCTCCAGAGAGTAAAGAGCTGATGAAACAATGGCAAGC 840
Db 1780 AAGGAAAGGCTCTTGAAGATCTCCAGAGAGTAAAGAGCTGATGAAACAATGGCAAGC 1839
QY 841 CTCGAGGTGAATTTGAAGCTCACACAGATGTTATTCACAACCTGATGAAACAGCCAA 900
Db 1840 CTCGAGGTGAATTTGAAGCTCACACAGATGTTATTCACAACCTGATGAAACAGCCAA 1899
QY 901 AAAATCTGAGATCCCTGAGAGTTCGATGATGATGATGATGATGATGATGATGATGATG 960
Db 1900 AAAATCTGAGATCCCTGAGAGTTCGATGATGATGATGATGATGATGATGATGATGATG 1959
QY 961 AACATGATCTCAATGAGAGTGAACCTTGGGAAAGTCTCAACATTAAGTCCCATTTG 1020
Db 1960 AACATGATCTCAATGAGAGTGAACCTTGGGAAAGTCTCAACATTAAGTCCCATTTG 2019
QY 1021 GAAGCAGTCTGACAGAGTGAAGGCTGACCTTTCTCTGACAGAACTTCTGATGAG 1080
Db 2020 GAAGCAGTCTGACAGAGTGAAGGCTGACCTTTCTCTGACAGAACTTCTGATGAG 2079
QY 1081 CTACAGCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 2080 CTACAGCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2139
QY 1141 GTTCAGAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Db 2140 GTTCAGAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2199
QY 1201 GTATATCATGATCTTGAAGCTGTAGAGATTTTCTGACAGAGCAGCTTTGAGAGGA 1260
Db 2200 GTATATCATGATCTTGAAGCTGTAGAGATTTTCTGACAGAGCAGCTTTGAGAGGA 2329
QY 1261 CTAGAGAACTCTACAGAGAGCCGAGAGAGTGCCTCTGAGAGAGAGCCGAGATGTC 1320

Db 2260 CTAGAGAACTCTACAGAGAGCCGAGAGAGTGCCTCTGAGAGAGAGCCGAGATGTC 2319
QY 1321 ACTGAGCTCTCAAGAAAGAGGCTGAGAGTCAATTAAGTGGGAAAAATTTGAACCTG 1380
Db 2320 ACTGAGCTCTCAAGAAAGAGGCTGAGAGTCAATTAAGTGGGAAAAATTTGAACCTG 2379
QY 1381 CACTCCGCTGACTGCGAGAGAAATATGATGAGACCTTTGAAAGACTCCAGAGACCTCAA 1440
Db 2380 CACTCCGCTGACTGCGAGAGAAATATGATGAGACCTTTGAAAGACTCCAGAGACCTCAA 2439
QY 1441 GAGGCCAGGATGAGTGAAGCTCAAGCTGCGCAGAGCTGAGGTGATCAAGGATCTCTG 1500
Db 2440 GAGGCCAGGATGAGTGAAGCTCAAGCTGCGCAGAGCTGAGGTGATCAAGGATCTCTG 2499
QY 1501 CAGCCGCTGAGGATCTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Db 2500 CAGCCGCTGAGGATCTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2559
QY 1561 CTTCGAGAGAAATTTGCGCTCTGAAAGAGAAAGTGAAGCAGTCAATGAGCTTCTGCG 1620
Db 2560 CTTCGAGAGAAATTTGCGCTCTGAAAGAGAAAGTGAAGCAGTCAATGAGCTTCTGCG 2619
QY 1621 CAGCTTACCACTTTGGGCAATTCAGCTCTCAAGCTTAACTTCAGCACTCTGAGAGACCTG 1680
Db 2620 CAGCTTACCACTTTGGGCAATTCAGCTCTCAAGCTTAACTTCAGCACTCTGAGAGACCTG 2679
QY 1681 AACACCAATGAGAGCTTTGCAAGTGGCTGAGAGACCAAGTCAAGGCACTGATGAA 1740
Db 2680 AACACCAATGAGAGCTTTGCAAGTGGCTGAGAGACCAAGTCAAGGCACTGATGAA 2739
QY 1741 GCCCAGAGGACCTTTGGTCAAGATCTGAGCACTTCTTCCAGCTGTCAGAGGCTCC 1800
Db 2740 GCCCAGAGGACCTTTGGTCAAGATCTGAGCACTTCTTCCAGCTGTCAGAGGCTCC 2799
QY 1801 TGGGAGAGGACCATCTCGCAAAACAAAGTCCCTACTATATCAACCAAGCACTCAACA 1860
Db 2800 TGGGAGAGGACCATCTCGCAAAACAAAGTCCCTACTATATCAACCAAGCACTCAACA 2859
QY 1861 ACTTGCTGAGGACCATCTCGCAAAACAAAGTCCCTACTATATCAACCAAGCACTCAACA 1920
Db 2860 ACTTGCTGAGGACCATCTCGCAAAACAAAGTCCCTACTATATCAACCAAGCACTCAACA 2919
QY 1921 GTCAAGTCTCAGCTTATGAGCTGCAAGTGAACCTCGAAGCTGCAAGAGCCCTTTC 1980
Db 2920 GTCAAGTCTCAGCTTATGAGCTGCAAGTGAACCTCGAAGCTGCAAGAGCCCTTTC 2979
QY 1981 TGGATCTCTGAGGCTGCAAGTGAACCTCGAAGCTGCAAGAGCCCTTTC 2040
Db 2980 TGGATCTCTGAGGCTGCAAGTGAACCTCGAAGCTGCAAGAGCCCTTTC 3039
QY 2041 AATGACCAAGCAGATGATCTGCAAGTGAACCTCGAAGCTGCAAGAGCCCTTTC 2100
Db 3040 AATGACCAAGCAGATGATCTGCAAGTGAACCTCGAAGCTGCAAGAGCCCTTTC 3099
QY 2101 CTGAGCAAGAGCAACAATTTGTCACAGTCTCTCTGCTGAGATGATGATGATGATGATG 2160
Db 3100 CTGAGCAAGAGCAACAATTTGTCACAGTCTCTCTGCTGAGATGATGATGATGATGATG 3159
QY 2161 TGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
Db 3160 TGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3219
QY 2221 ACTGAGCATATTTCCCTGTGAAGCAATTTGGAAGCAAGTACAGATCACTTTTCAAG 2280
Db 3220 ACTGAGCATATTTCCCTGTGAAGCAATTTGGAAGCAAGTACAGATCACTTTTCAAG 3279
QY 2281 CAGTGGCAAGTTCACAGAGATTTTGTGACAGAGCAGGCTGAGGCTCTCTGATGATG 2340
Db 3280 CAGTGGCAAGTTCACAGAGATTTTGTGACAGAGCAGGCTGAGGCTCTCTGATGATG 2399
QY 2341 TCTATCAAAATTCAGAGACAGTGGGTGAAGTTCATCTTTGGGGGAGTAACTTGAAG 2400
Db 3340 TCTATCAAAATTCAGAGACAGTGGGTGAAGTTCATCTTTGGGGGAGTAACTTGAAG 3399

QY 2401 C 2401
Db 3400 C 3400

RESULT 2

US-09-845-416-29
Sequence 29, Application US/09845416
Publication No. US2003017312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 29
LENGTH: 4825
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-29

Query Match 100.0%; Score 2401; DB 13; Length 4825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTTCATTGATGAGAGTGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1757 GGCAGTTCATTGATGAGAGTGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAA 1816
QY 61 GTATTATGCTGGCTTTCTTCTGCTGAGACACATTGCAAGCAAGAGATTCTAAT 120
Db 1817 GTATTATGCTGGCTTTCTTCTGCTGAGACACATTGCAAGCAAGAGATTCTAAT 1876
QY 121 GATGTGAAGTGTGGAAGAACCACTTCTACTCATGAGGGGTACATGATGATTTGACA 180
Db 1877 GATGTGAAGTGTGGAAGAACCACTTCTACTCATGAGGGGTACATGATGATTTGACA 1936
QY 181 GCCCATCAGGSCCGGGTGTGAATATTTCTCAATGAGGAACTGAATGGAACAGGA 240
Db 1937 GCCCATCAGGSCCGGGTGTGAATATTTCTCAATGAGGAACTGAATGGAACAGGA 1996
QY 241 AAATTATCAGAAATGAAAGAACTGAAGTCAAGAGAGATGATCTCTTAATTGACA 300
Db 1997 AAATTATCAGAAATGAAAGAACTGAAGTCAAGAGAGATGATCTCTTAATTGACA 2056
QY 301 TGGGAATGCTCAGGGTAGTACAGTGAAGAAACAAAGCAATTTACATAGATTTAATG 360
Db 2057 TGGGAATGCTCAGGGTAGTACAGTGAAGAAACAAAGCAATTTACATAGATTTAATG 2116
QY 361 GATCTCCGAATTCGAAGAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAA 420
Db 2117 GATCTCCGAATTCGAAGAACTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAA 2176
QY 421 ACAAGAAATGAGAGAAAGAGCTCTTGAACCTGATCTTGAAGACCTTAAGAGCCAGTA 480
Db 2177 ACAAGAAATGAGAGAAAGAGCTCTTGAACCTGATCTTGAAGACCTTAAGAGCCAGTA 2236
QY 481 CAACACATTAAGTGTCTTCAAGAAATGATAGACAAAGCAAGTCAAGGCTCAATCTCTC 540
Db 2237 CAACACATTAAGTGTCTTCAAGAAATGATAGACAAAGCAAGTCAAGGCTCAATCTCTC 2296
QY 541 ACTCAATGAGTGTGTGATGATGATCTAGTGAAGTCAAGCAAGTCAAGGCTTTGAA 600
Db 2297 ACTCAATGAGTGTGTGATGATGATCTAGTGAAGTCAAGCAAGTCAAGGCTTTGAA 2356
QY 601 GAACAACTTAAGGATTTGGAGATGATGGGCAAACTCTGTAGATGACAGAGACCGC 660

Db 2357 GAACAACTTAAGGATTTGGAGATGATGGGCAAACTCTGTAGATGACAGAGACCGC 2416
QY 661 TGGGTTCTTTTCAAGACACTCATAGATTACTGCAACAGTTCCCTGGACCTGAAAG 720
Db 2417 TGGGTTCTTTTCAAGACACTCATAGATTACTGCAACAGTTCCCTGGACCTGAAAG 2476
QY 721 TTTGCTGCTGGCTTACAGAACTGAAACAACTGCAATGCTCAAGAGTCTCCCT 780
Db 2477 TTTGCTGCTGGCTTACAGAACTGAAACAACTGCAATGCTCAAGAGTCTCCCT 2536
QY 781 AAGGAAAGGCTCTTGAAGACTCCAGAGAGTAAAGAGCTGATGAAACAAATGGCAAGC 840
Db 2537 AAGGAAAGGCTCTTGAAGACTCCAGAGAGTAAAGAGCTGATGAAACAAATGGCAAGC 2596
QY 841 CTCGAAGTGAATTTGAAGCTTACACAGATTTTATCACAACTTGATGAAACAGCCAA 900
Db 2597 CTCGAAGTGAATTTGAAGCTTACACAGATTTTATCACAACTTGATGAAACAGCCAA 2656
QY 901 AAAATCCTGAGATCCTGGAAAGTCCGATGATGCAATCTGTACAAAGAGCTTGGAT 960
Db 2657 AAAATCCTGAGATCCTGGAAAGTCCGATGATGCAATCTGTACAAAGAGCTTGGAT 2716
QY 961 AACTGAACCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020
Db 2717 AACTGAACCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2776
QY 1021 GAAGCAGTCTGACACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1080
Db 2777 GAAGCAGTCTGACACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2836
QY 1081 CTACAGCTGAAGATGATGAATTAAGCCGACAGCACTTATGAGGCACTTTGACACA 1140
Db 2837 CTACAGCTGAAGATGATGAATTAAGCCGACAGCACTTATGAGGCACTTTGACACA 2896
QY 1141 GTTCAGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 2897 GTTCAGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2956
QY 1201 GTAATCAGATGATCTTGAAGCTGATGAATATTTCTGACAGAGCACTTTGGAAGA 1260
Db 2957 GTAATCAGATGATCTTGAAGCTGATGAATATTTCTGACAGAGCACTTTGGAAGA 3016
QY 1261 CTAGAGAACTTACAGAGAGCCAGAGAGTCTCTGAGAGAGAGCCAGATGTC 1320
Db 3017 CTAGAGAACTTACAGAGAGCCAGAGAGTCTCTGAGAGAGAGCCAGATGTC 3076
QY 1321 ACTCGGCTCTACAGAAAGAGCTGAGAGTCAATCTGAGTGGGAAAAATTGAACCTG 1380
Db 3077 ACTCGGCTCTACAGAAAGAGCTGAGAGTCAATCTGAGTGGGAAAAATTGAACCTG 3136
QY 1381 CACTCGCTGACTGGCAGAGAAATTAATGATGAGACCTTGAAGAGCTCAGAGACTTCA 1440
Db 3137 CACTCGCTGACTGGCAGAGAAATTAATGATGAGACCTTGAAGAGCTCAGAGACTTCA 3196
QY 1441 GAGGCGAGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500
Db 3197 GAGGCGAGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3256
QY 1501 CAGCCGTTGGGAGATCTCTCATTAATGACTCTCTCAAGATCACTTGAAGAACTCAAGCA 1560
Db 3257 CAGCCGTTGGGAGATCTCTCATTAATGACTCTCTCAAGATCACTTGAAGAACTCAAGCA 3316
QY 1561 CTTGAGAGAAATTTGGCTCTTGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1620
Db 3317 CTTGAGAGAAATTTGGCTCTTGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3376
QY 1621 CAGCTTACCACTTTGGGATTCAGAGTCAACCGTAACTCAAGCACTCTGGAAGACCTG 1680
Db 3377 CAGCTTACCACTTTGGGATTCAGAGTCAACCGTAACTCAAGCACTCTGGAAGACCTG 3436
QY 1681 AAACCAAGTGAAGCTTCTGAGAGTGCCTGAGAGACCGAGTCAAGCAGCTGATGAA 1740
Db 3437 AAACCAAGTGAAGCTTCTGAGAGTGCCTGAGAGACCGAGTCAAGCAGCTGATGAA 3496

1741 GCCACAGAGGACTTTGGTCAGCATCTCAGACATTTCTTCCAGTCTGTCCAGGGTCCC 1800
1801 TGGGAGAGAGCCATCTGGCCAAAGAGTGGCCCTACTATATCAACCCAGAGATCAAAAC 1860
1861 ACTTGTGGAGACCAATCCAAATAGACAGAGCTTACAGCTTTAGCTGCTGATTAAT 1920
1921 GTGAGATTCCTGAGCTTATAGAGCTCCAGTAACTCCGAAAGCTGCAAGAGCCCTTGGC 1980
1981 TTGATCTCTTGGAGCTGTGAGCTGAGCTGATGATGATGATGATGATGATGATGATGAT 2040
2041 AATGACAGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
2101 CTGAGCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2160
2161 TGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
2221 ACTGCAATCATTTCTGTGTGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2280
2281 CAAGTGGAGGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2340
2341 TCTATCCAAATTCAGAGCAGTGGGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2400
2401 C 2401
4157 C 4157

RESULT 3
US-09-845-416-35
Sequence 35, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DEL142
CURRENT APPLICATION NUMBER: US/09/845,416
PRIOR FILING DATE: 2001-04-30
PRIORITY FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 35
LENGTH: 4848
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 100.0%; Score 2401; DB 13; Length 4848;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGCAGTTCATGATGAGAGAGTGAAGTAACTGGACCGTATCAACAGCTTTAGAGAA 60
1780 GGCAGTTCATGATGAGAGAGTGAAGTAACTGGACCGTATCAACAGCTTTAGAGAA 1839
61 GTATTATCTGTGGCTTTCTTCTGTGAGAGACATTTGCAAGCAAGAGAGATTTCTAAT 120
1840 GTATTATCTGTGGCTTTCTTCTGTGAGAGACATTTGCAAGCAAGAGAGATTTCTAAT 1899
121 GATGGAAGTGTGAAGAACCACTTCACTCATGAGGGGTACATGATGATTTGACA 180
1900 GATGGAAGTGTGAAGAACCACTTCACTCATGAGGGGTACATGATGATTTGACA 1959
181 GCCCATCAGGGCCGGGTGGTAAATTTCTCAATTTGGAAGTAACTGATTTGACA 240
1960 GCCCATCAGGGCCGGGTGGTAAATTTCTCAATTTGGAAGTAACTGATTTGACA 2019
241 AAATTATGAAAGATGAAAGAACTGAAGTACAGAGCAGATGATCTCTAAATTCAGA 300
2020 AAATTATGAAAGATGAAAGAACTGAAGTACAGAGCAGATGATCTCTAAATTCAGA 2079
301 TGGGATGCTCAGGGTGTAGTGAATGAGAAAAAAGCAATTTACATGAGTTTAAATG 360
2080 TGGGATGCTCAGGGTGTAGTGAATGAGAAAAAAGCAATTTACATGAGTTTAAATG 2139
361 GATCTCCAGAAATCAGAAATCTGAAGAGTTGATGATGCTGCTTAAAGAAAGAGAA 420
2140 GATCTCCAGAAATCAGAAATCTGAAGAGTTGATGATGCTGCTTAAAGAGAAAGAA 2199
421 ACAAGGAAATGAGAGAGAGCTCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 480
2200 ACAAGGAAATGAGAGAGAGCTCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 2259
481 CAACAACTATGAGTCTTCAAGAGATCTGAACAGAAAGAGTCAAGTCAATTTCTCTC 540
2260 CAACAACTATGAGTCTTCAAGAGATCTGAACAGAAAGAGTCAAGTCAATTTCTCTC 2319
541 ACTCAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
2320 ACTCAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2379
601 GAACAACTTATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
2380 GAACAACTTATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2439
661 TGGTTCCTTTTCAAGACATCTATGATTAATCTGCAACAGTCCCTGGACCTGGAAAG 720
2440 TGGTTCCTTTTCAAGACATCTATGATTAATCTGCAACAGTCCCTGGACCTGGAAAG 2499
721 TTTCTTCTGCTGCTTACAGAGCTGAAACAATGCTCAATGCTTACAGAGTGTACCCGT 780
2500 TTTCTTCTGCTGCTTACAGAGCTGAAACAATGCTCAATGCTTACAGAGTGTACCCGT 2559
781 AAGGAAAGGCTCTTGAAGACCTCCAGAGAGTAAAGAGCTGATTAATGGAAGAAC 840
2560 AAGGAAAGGCTCTTGAAGACCTCCAGAGAGTAAAGAGCTGATTAATGGAAGAAC 2619
841 CTCGAAGTGAATGAGAGCTCAGACAGATGTTTATCAACCTGGATGAAACAGCCAA 900
2620 CTCGAAGTGAATGAGAGCTCAGACAGATGTTTATCAACCTGGATGAAACAGCCAA 2679
901 AAAATCTGAGATCCCTGGAAGGTTCCGATGATGAGAGTCTGTTTCAAAAGAGCTTTGAT 960
2680 AAAATCTGAGATCCCTGGAAGGTTCCGATGATGAGAGTCTGTTTCAAAAGAGCTTTGAT 2739
961 AACATGAATTCAGAGAGAGTGAACCTGGAAGAAAGTCTCTCAACATTAAGTCCCATTTG 1020
2740 AACATGAATTCAGAGAGAGTGAACCTGGAAGAAAGTCTCTCAACATTAAGTCCCATTTG 2799
1021 GAAGCAGTCTGACAGAGTGAAGAGCTTGAACCTTTCTGACAGAACTTGTGATGTGG 1080
2800 GAAGCAGTCTGACAGAGTGAAGAGCTTGAACCTTTCTGACAGAACTTGTGATGTGG 2859

QY	1081	CTACAGCTGAAAGATGATGAAATTAAAGCCGAGGACACCTAATTGAGGGACATTTCACGCA	1140
Dp	2860	CTACAGCTGAAAGATGATGAAATTAAAGCCGAGGACACCTAATTGAGGGACATTTCACGCA	2919
QY	1141	GTTTCAAGAACGAACGATGTATCATTAAGGCGCTTTCAGAGGAATTGAAACTTAAGAACCT	1200
Dp	2920	GTTTCAAGAACGAACGATGTATCATTAAGGCGCTTTCAGAGGAATTGAAACTTAAGAACCT	2979
QY	1201	GTAATCATGATTAAGTACTCTTGAAGACTGTACGAAATATTTTCGACAGAGCAGCCCTTGTGAAGA	1260
Dp	2980	GTAATCATGATTAAGTACTCTTGAAGACTGTACGAAATATTTTCGACAGAGCAGCCCTTGTGAAGA	3039
QY	1261	CTAAGAGAACTCTTACACAGAGACCCAGAGAGCTGCTCTCTGAGAGAGAGCCAGATATGTC	1320
Dp	3040	CTAAGAGAACTCTTACACAGAGACCCAGAGAGAGCTGCTCTCTGAGAGAGAGCCAGATATGTC	3099
QY	1321	ACTCGGCTTCTACCAAAAGCAGGCTGAGAGAGTCAATACTGATGGAGAAAATTTGAACCTG	1380
Dp	3100	ACTCGGCTTCTACCAAAAGCAGGCTGAGAGAGTCAATACTGATGGAGAAAATTTGAACCTG	3159
QY	1381	CACATCCGTGACTGGCCAGAGAAAATAGATGACACCTTGAAAGATTCACAGAACTTCA	1440
Dp	3160	CACATCCGTGACTGGCCAGAGAAAATAGATGACACCTTGAAAGATTCACAGAACTTCA	3219
QY	1441	GAGGCCACGATGAGCTGGAACCTCAAGCTGGCCAAAGCTGAGTGAATCAAGGACCTCTGG	1500
Dp	3220	GAGGCCACGATGAGCTGGAACCTCAAGCTGGCCAAAGCTGAGTGAATCAAGGACCTCTGG	3279
QY	1501	CAGCCCGTGGGCGGATCTCTCATTTGACTCTTCCAAAGTACCTTCGAGAAAGTCAAGGCA	1560
Dp	3280	CAGCCCGTGGGCGGATCTCTCATTTGACTCTTCCAAAGTACCTTCGAGAAAGTCAAGGCA	3339
QY	1561	CTTGGAGAGAGAAATTGGCGCTCTGAAAAGAGAACTGAGCCACGTCAAATGACCTTGCTGC	1620
Dp	3340	CTTGGAGAGAGAAATTGGCGCTCTGAAAAGAGAACTGAGCCACGTCAAATGACCTTGCTGC	3399
QY	1621	CAGCTTACCACTTTGGGGCATTCAGTCTCAACCGTAATACTTCAGACACTTCGAAAGACTG	1680
Dp	3400	CAGCTTACCACTTTGGGGCATTCAGTCTCAACCGTAATACTTCAGACACTTCGAAAGACTG	3459
QY	1681	AACACCAAGATGAAGCTTCTGACAGTGGCCCTGCAGAGACCAAGTCAGGACAGCTCATGAA	1740
Dp	3460	AACACCAAGATGAAGCTTCTGACAGTGGCCCTGCAGAGACCAAGTCAGGACAGCTCATGAA	3519
QY	1741	GCCCAACAGGACCTTTGGTCCAGCATCTTCAGACATTTCTTCCAGTCTGTCCAGGATCCC	1800
Dp	3520	GCCCAACAGGACCTTTGGTCCAGCATCTTCAGACATTTCTTCCAGTCTGTGTCCAGGATCCC	3579
QY	1801	TGGGAGAGAGCATCTCGCCAAACAAAGTGCCTCACTATATCAACCAACAGACCTCAACA	1860
Dp	3580	TGGGAGAGAGCATCTCGCCAAACAAAGTGCCTCACTATATCAACCAACAGACCTCAACA	3639
QY	1861	ACTTGCTGGGACCAATCCCAAAATGACAGAGCTTCAACGATCTTTAGCTGACCTGAATAT	1920
Dp	3640	ACTTGCTGGGACCAATCCCAAAATGACAGAGCTTCAACGATCTTTAGCTGACCTGAATAT	3699
QY	1921	GTCAGATTCTAGCTTATAGGACTGCCATGAACTCCGAAAGCTCGAAGACTCGAAGAGCCCTTTGC	1980
Dp	3700	GTCAGATTCTAGCTTATAGGACTGCCATGAACTCCGAAAGCTCGAAGAGACTCGAAGAGCCCTTTGC	3759
QY	1981	TTGGATCTCTTGAAGCTGTCAAGCTGCAATGTATGCTTTGGACCAAGCAACCTCAAGCA	2040
Dp	3760	TTGGATCTCTTGAAGCTGTCAAGCTGCAATGTATGCTTTGGACCAAGCAACCTCAAGCA	3819
QY	2041	AATGACCAAGCCATGATATCTGACAGATTATATTTGTTTGACCACTAATTTATGACGCG	2100
Dp	3820	AATGACCAAGCCATGATATCTGACAGATTATATTTGTTTGACCACTAATTTATGACGCG	3879
QY	2101	CTGAGAGCAAGCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATATATGTGCTGAC	2160
Dp	3880	CTGAGAGCAAGCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATATATGTGCTGAC	3939

Oy	2161	TGCGTGGCAATGTTTATGTAAACGGAGCAACAAGAGAGATCCCTGTCTCTTTAAA	2280
Db	3940	TGGCTGCTGAATGTTTATGTAAACGGAGCAACAAGAGAGATCCCTGTCTCTTTAAA	3999
Oy	2221	ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAGAACAATAACATTTCGAAG	2280
Db	4000	ACTGGCATCATTTCCCTGTGTAAAGCACCATTGGAAACAAGTAACAATATCCCTTTGAAG	4059
Oy	2281	CAACTGGCAAGTTCACAGGATTTTGACACAGGCAGGCTGGGCTCTCTCTGCATGAT	2340
Db	4060	CMACTGGCAAGTTCACAGATTTTGTGACCAAGGCAGGCTGGGCTCTCTCTGCATGAT	4119
Oy	2341	TCATATCCAAATTCACAGACAGTGGGTGAAGTTGCATCCTTTGGGGGACAGTAACATTGAG	2400
Db	4120	TCATATCCAAATTCACAGACAGTGGGTGAAGTTGCATCCTTTGGGGGACAGTAACATTGAG	4179
Oy	2401	C 2401	
Db	4180	C 4180	

```

RESULT 4
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication NO. US2003017312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845, 416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match          100.0%; Score 2401; DB 13; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGCAGTTCATTGATGAGAGCTGAAGTAACCTGGACCGTTATCAACAAGCTTTGAGAA 60
DB      1592  GGCAGTTCATTGATGAGAGCTGAAGTAACCTGGACCGTTATCAACAAGCTTTGAGAA 20511

QY      61  GTATTATCGGGCTCTCTTCGCGTGAAGACACATGGCAAGACAAGAGAGATTCTTAAT 120
DB      2052  GTATTATCGGGCTCTCTTCGCGTGAAGACACATGGCAAGACAAGAGAGATTCTTAAT 21111

QY      121  GATGTGGAAGTGTGTGAAGAACAAGTTTCATACCTCATGAGGGGTACATGATGATTTGACA 180
DB      2112  GATGTGGAAGTGTGTGAAGAACAAGTTTCATACCTCATGAGGGGTACATGATGATTTGACA 21711

QY      181  GCCCATCAGAGGCCGGGGTGTGTAATATCTTCAATTTGGGAAGTAGCTGATTGGAACAGA 240
DB      2172  GCCCATCAGAGGCCGGGGTGTGTAATATCTTCAATTTGGGAAGTAGCTGATTGGAACAGA 22311

QY      241  AAATTATCAGAGATGAGAAGAACTGAAGTCAAGAGACAGATGATCTCTTAATTTACAGA 300
DB      2232  AAATTATCAGAGATGAGAAGAACTGAAGTCAAGAGACAGATGATCTCTTAATTTACAGA 22911

QY      301  TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAGCAATTTACATAGATTTTAATG 360
DB      2292  TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAGCAATTTACATAGATTTTAATG 23511

QY      361  GATGTCCAGATCCGAAGTGAAGAAAGATTGAATCACTGGCTACAAACAAACAGAGAGAAGA 420
DB      2352  GATGTCCAGATCCGAAGTGAAGAAAGATTGAATCACTGGCTACAAACAAACAGAGAGAAGA 241111

```

QY 421 ACAAGAAATGAGAGAGAGCTCTTGACCTGATCTTGAGAGACTTAAACGCGCAAGTA 480
Db 2412 ACAAGAAATGAGAGAGAGCTCTTGACCTGATCTTGAGAGACTTAAACGCGCAAGTA 2471
QY 481 CAACACATTAAGGTGCTTCAAGAGATCTAGAACAAAGACAGGTGCTGCTGCTC 540
Db 2472 CAACACATTAAGGTGCTTCAAGAGATCTAGAACAAAGACAGGTGCTGCTGCTC 2531
QY 541 ACTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 2532 ACTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2591
QY 601 GAAACATTAAGGTGCTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 2592 GAAACATTAAGGTGCTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2651
QY 661 TGGGTCTTTTACAAAGACATCATAGATTAATGACAACTGCTGCTGCTGCTGCTGCTGCT 720
Db 2652 TGGGTCTTTTACAAAGACATCATAGATTAATGACAACTGCTGCTGCTGCTGCTGCTGCT 2711
QY 721 TTTCTGCTGCTGCTTACAAAGCTGAAACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 2712 TTTCTGCTGCTGCTTACAAAGCTGAAACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2771
QY 781 AAGGAAAGGCTCTAGAGAGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 2772 AAGGAAAGGCTCTAGAGAGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2831
QY 841 CTCGAAAGGTGAAATGAAAGCTGACACAGATGTTATGACAACTGATGAGAGAGAGAG 900
Db 2832 CTCGAAAGGTGAAATGAAAGCTGACACAGATGTTATGACAACTGATGAGAGAGAGAG 2891
QY 901 AAAATCTGAGATCCCTGAGAGGTTCGATGATGATGATGATGATGATGATGATGATGAT 960
Db 2892 AAAATCTGAGATCCCTGAGAGGTTCGATGATGATGATGATGATGATGATGATGATGAT 2951
QY 961 AACATGACTCAAGTGAAGTGAAGTTCGAAAGAGTCTGCAACTTAGGCTCCATTG 1020
Db 2952 AACATGACTCAAGTGAAGTGAAGTTCGAAAGAGTCTGCAACTTAGGCTCCATTG 3011
QY 1021 GAAAGCAATGCTGACCAAGTGAAGGTCTGCACTTCTCTGCAAGAACTTCTGCTGCTG 1080
Db 3012 GAAAGCAATGCTGACCAAGTGAAGGTCTGCACTTCTCTGCAAGAACTTCTGCTGCTG 3071
QY 1081 CTACAGCTGAAAGATGATGATTAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 3072 CTACAGCTGAAAGATGATGATTAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3131
QY 1141 GTTCAGAGCAGAAAGATGATCATAGGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 3132 GTTCAGAGCAGAAAGATGATCATAGGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAG 3191
QY 1201 GAAATCATGAGTACTCTGAGAGCTGATGAGATTTTCTGACAGAGAGAGAGAGAGAG 1260
Db 3192 GAAATCATGAGTACTCTGAGAGCTGATGAGATTTTCTGACAGAGAGAGAGAGAGAG 3251
QY 1261 CTAGAGAACTCTACAG 1320
Db 3252 CTAGAGAACTCTACAG 3311
QY 1321 ACTGGGCTTTCTAG 1380
Db 3312 ACTGGGCTTTCTAG 3371
QY 1381 CACTCGCTGATGCTGAG 1440
Db 3372 CACTCGCTGATGCTGAG 3431
QY 1441 GAGGCCAGGATGAG 1500
Db 3432 GAGGCCAGGATGAG 3491

QY 1501 CAGCCCGTGGGAGATCTCTCTCAATGATCTCTCAAGATCACTTGAAGAGAGAGAGAG 1560
Db 3492 CAGCCCGTGGGAGATCTCTCTCAATGATCTCTCAAGATCACTTGAAGAGAGAGAGAG 3551
QY 1561 CTTGAG 1620
Db 3552 CTTGAG 3611
QY 1621 CAGTTACCATTTGGGAGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 3612 CAGTTACCATTTGGGAGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 3671
QY 1681 AACACAGATGAG 1740
Db 3672 AACACAGATGAG 3731
QY 1741 GCCCAG 1800
Db 3732 GCCCAG 3791
QY 1801 TGGAG 1860
Db 3792 TGGAG 3851
QY 1861 ACTTCTGAG 1920
Db 3852 ACTTCTGAG 3911
QY 1921 GTGAGATTTCTGAG 1980
Db 3912 GTGAGATTTCTGAG 3971
QY 1981 TTGAGATTTCTGAG 2040
Db 3972 TTGAGATTTCTGAG 4031
QY 2041 AATGACAG 2100
Db 4032 AATGACAG 4091
QY 2101 CTGAG 2160
Db 4092 CTGAG 4151
QY 2161 TGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db 4152 TGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4211
QY 2221 ACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 4212 ACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4271
QY 2281 CAAGTGGCAAGTTCAG 2340
Db 4272 CAAGTGGCAAGTTCAG 4331
QY 2341 TCTATCAAAATTCAG 2400
Db 4332 TCTATCAAAATTCAG 4391
QY 2401 C 2401
Db 4392 C 4392

RESULT 5
US-09-845-416-6
; Sequence 5, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THERIOF

```

FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 3999
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-6

```

```

Query Match      93.7%; Score 2250; DB 13; Length 3999;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

```

```

QY 1 GGCAGTTCAATTGAGAGAGTAAACCTGACCGTTATCAACAGCTTTAGAGAA 60
DB 1000 GGCAGTTCAATTGAGAGAGTAAACCTGACCGTTATCAACAGCTTTAGAGAA 1059
QY 61 GTATTATCGTGCTTTCTGCTGAGACACATTGCAAGCAAGAGAGATTCTAT 120
DB 1060 GTATTATCGTGCTTTCTGCTGAGACACATTGCAAGCAAGAGAGATTCTAT 1119
QY 121 GATGTGAGAGTGTGAAACCACTTCTATCTCATGAGGGGTATGATGATTTGACA 180
DB 1120 GATGTGAGAGTGTGAAACCACTTCTATCTCATGAGGGGTATGATGATTTGACA 1179
QY 181 GCCCATCAGGCGCGGTTGGTAAATTCTACATTGGGAGTAACTGATTTGAGACAGA 240
DB 1180 GCCCATCAGGCGCGGTTGGTAAATTCTACATTGGGAGTAACTGATTTGAGACAGA 1239
QY 241 AAATTATCGAAGATGAAGAACTGAGTACAGAGCAGATGATCTCTTAATTCAAGA 300
DB 1240 AAATTATCGAAGATGAAGAACTGAGTACAGAGCAGATGATGATCTCTTAATTCAAGA 1239
QY 301 TGGGAAATGCTCAGGCTAGCTAGCATGAGAAACAAAGCAATTTTCATGAGTTTATG 360
DB 1300 TGGGAAATGCTCAGGCTAGCTAGCATGAGAAACAAAGCAATTTTCATGAGTTTATG 1359
QY 361 GATCTTCAGAAATCAGAACTGAGAGAGTGAATGATGCTGCTAACAAAAACAGAGAGAA 420
DB 1360 GATCTTCAGAAATCAGAACTGAGAGAGTGAATGATGCTGCTAACAAAAACAGAGAGAA 1419
QY 421 ACAAAGAAATGAGAGAGAGAGCTCTTGAGACTGATCTTGAAGACTTAAACGCCAAGTA 480
DB 1420 ACAAAGAAATGAGAGAGAGAGCTCTTGAGACTGATCTTGAAGACTTAAACGCCAAGTA 1479
QY 481 CAACAACATAGGTGCTTCAGAGAGATAGAGAAAGAAACAAGTCAGGGTCAATTTCTC 540
DB 1480 CAACAACATAGGTGCTTCAGAGAGATAGAGAAAGAAACAAGTCAGGGTCAATTTCTC 1539
QY 541 ACTCACATGTGTGTGTAGTATGATGATTAATGATGAGATCACGCACTGCTGCTTTGAA 600
DB 1540 ACTCACATGTGTGTGTAGTATGATGATTAATGATGAGATCACGCACTGCTGCTTTGAA 1539
QY 601 GAACAACCTAAGGTATGAGAGATGAGTGGCAAAACATCTGATGATGAGAGAGAGAGAGC 660
DB 1600 GAACAACCTAAGGTATGAGAGATGAGTGGCAAAACATCTGATGATGAGAGAGAGAGC 1659
QY 661 TGGGTTCTTTTCAAGAC----- 678
DB 1660 TGGGTTCTTTTCAAGACAGAGCTGACCTGACCTGAGCTGACCACTATTGAGAGCTCTCT 1719
QY 679 ----- 678
DB 1720 CCTACTCAGACTGTACTCTGTGTACACAACTGTGTACTAGAGAAACTGCCATCTCC 1779
QY 679 -----ACTCATAGATTACTGCAACAG 699
DB 1780 AAACAGAAATGACATCTTCTCTTGAATGAGAGTACCTACATGATTAATCTGAAACAG 1839

```

```

QY 700 TTCCCTGAGACCTGGAAAAGTTTCTTGCTGCTTACAGAGCTGAAAACACTGCCAAT 759
DB 1840 TTCCCTGAGACCTGGAAAAGTTTCTTGCTGCTTACAGAGCTGAAAACACTGCCAAT 1899
QY 760 GTCTTACAGATGCTACCCCTTAGAGAAAGGCTCTTGAAGACTCCACAGAGAGTAAAGAG 819
DB 1900 GTCTTACAGATGCTACCCCTTAGAGAAAGGCTCTTGAAGACTCCACAGAGAGTAAAGAG 1959
QY 820 CTGATGAAACAATGAGCAAGCTCCAGGTGAATTAAGCTCACAGATGTTATAC 879
DB 1960 CTGATGAAACAATGAGCAAGCTCCAGGTGAATTAAGCTCACAGATGTTATAC 2019
QY 880 AACTGATGAAAACAGCCAAAATCTGAGATCCCTGAGAGTTCCGATGATGCAATC 939
DB 2020 AACTGATGAAAACAGCCAAAATCTGAGATCCCTGAGAGTTCCGATGATGCAATC 2079
QY 940 CTGTTCAAGACGTTTGAATACATGAACTCAAGTGAAGTGAATCTGAGAAAAGTCT 999
DB 2080 CTGTTCAAGACGTTTGAATACATGAACTCAAGTGAAGTGAATCTGAGAAAAGTCT 2139
QY 1000 CTCAACATTAGGTCCTCAATTTGAAAGCAAGTCTGACCAAGTGAAGGCTGCACTTTCT 1059
DB 2140 CTCAACATTAGGTCCTCAATTTGAAAGCAAGTCTGACCAAGTGAAGGCTGCACTTTCT 2199
QY 1060 CTGAGAGAACTTCTGTGTGTGTGCTACAGCTGAAGAGATGAATTAAGCCGAGAGCACT 1119
DB 2200 CTGAGAGAACTTCTGTGTGTGTGCTACAGCTGAAGAGATGAATTAAGCCGAGAGCACT 2259
QY 1120 ATTGAGGCGCACTTTCAGAGCTTCAAGAGCAGAGATGATCATGAGGCTTCAAGAG 1179
DB 2260 ATTGAGGCGCACTTTCAGAGCTTCAAGAGCAGAGATGATCATGAGGCTTCAAGAG 2319
QY 1180 GAATTGAAAATTAAGAACTGTAATGATGATGATCTTTGAGACTGATGATATTTCTG 1239
DB 2320 GAATTGAAAATTAAGAACTGTAATGATGATGATCTTTGAGACTGATGATATTTCTG 2379
QY 1240 ACAAGAGAGCTTGTGAGAGAGCTAGAGAACTGTACAGAGCCAGAGAGCTGCTCT 1299
DB 2380 ACAAGAGAGCTTGTGAGAGAGCTAGAGAACTGTACAGAGCCAGAGAGCTGCTCT 2439
QY 1300 GAGAGAGAGCCAGAGATGCTCACTGCTTCTTCAAGAGCAGCTGAGAGAGTCAATACT 1359
DB 2440 GAGAGAGAGCCAGAGATGCTCACTGCTTCTTCAAGAGCAGCTGAGAGAGTCAATACT 2499
QY 1360 GAGTGGGAAAATTAAGAACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1419
DB 2500 GAGTGGGAAAATTAAGAACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2559
QY 1420 GAAAGACTCCAGAGACTTCAAGAGCCAGAGATGAGCTGAGCTCAGCTGCGCCAGCT 1479
DB 2560 GAAAGACTCCAGAGACTTCAAGAGCCAGAGATGAGCTGAGCTCAGCTGCGCCAGCT 2619
QY 1480 GAGGTATCAAGGATCTCTGAGAGCCGCTGAGCTGCTCTTCAATGATCTCTCCAAAT 1539
DB 2620 GAGGTATCAAGGATCTCTGAGAGCCGCTGAGCTGCTCTTCAATGATCTCTCCAAAT 2679
QY 1540 CACCTCGAAGAAAGTCAAGGCACTTCAAGAGAAATTTGCGCTCTGAAAGAGAAAGTGAAG 1599
DB 2680 CACCTCGAAGAAAGTCAAGGCACTTCAAGAGAAATTTGCGCTCTGAAAGAGAAAGTGAAG 2739
QY 1600 CACCTCAATACCTTCTGCGCAAGCTTCAACCTTTGAGGCACTTCAACCTGATTAAC 1659
DB 2740 CACCTCAATACCTTCTGCGCAAGCTTCAACCTTTGAGGCACTTCAACCTGATTAAC 2799
QY 1660 CTGAGACTCTGGAAGACCTGAAACACCAATGGAAGCTTCTGAGGTGCGCTGAGAGAC 1719
DB 2800 CTGAGACTCTGGAAGACCTGAAACACCAATGGAAGCTTCTGAGGTGCGCTGAGAGAC 2859
QY 1720 CGAGTCAAGCAAGCTGATGAAGCCACAGAGACTTTGATCAAGATCTTCAAGCACTTTCT 1779
DB 2860 CGAGTCAAGCAAGCTGATGAAGCCACAGAGACTTTGATCAAGATCTTCAAGCACTTTCT 2919
QY 1780 TCCAGCTGTGTCAAGGTCCTGAGAGAGAGCAATCTGCGCAAAACAAATGCGCTACTAT 1839

```

Db 2920 TCCAGTGTCTCCAGGATCCCTGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTAT 2979
Qy 1840 ATCAACCAAGAGACTCAAACTACTGCTGGAGACCATCCCAAAATGACAGAGCTTACAG 1899
Db 2980 ATCAACCAAGAGACTCAAACTACTGCTGGAGACCATCCCAAAATGACAGAGCTTACAG 3039
Qy 1900 TCTTTAGCTGACCTGAATATATGTCAGATTCTCAAGCTTTAGAGCTGCCATGAAACTCCGA 1959
Db 3040 TCTTTAGCTGACCTGAATATATGTCAGATTCTCAAGCTTTAGAGCTGCCATGAAACTCCGA 3099
Qy 1960 AGACTGCAAGAGCCCTTGTGCTGATCTCTGAGCCCTGAGCTGATGATGAGCTTGG 2019
Db 3100 AGACTGCAAGAGCCCTTGTGCTGATCTCTGAGCCCTGAGCTGATGATGAGCTTGG 3159
Qy 2020 GACCAACAACCTCAAGCAAAATGACAGCCCATGATATCTGCAAGATTATTAATGT 2079
Db 3160 GACCAACAACCTCAAGCAAAATGACAGCCCATGATATCTGCAAGATTATTAATGT 3219
Qy 2080 TTGACCACTATTATGACCGCTGAGAGCAAGCAAAATTTGCTCAAGCTTCTC 2139
Db 3220 TTGACCACTATTATGACCGCTGAGAGCAAGCAAAATTTGCTCAAGCTTCTC 3279
Qy 2140 TCCGTGATATGTCGTAACCTGCTGATGTTATGATACGAGAGCAAGAGAG 2199
Db 3280 TCCGTGATATGTCGTAACCTGCTGATGTTATGATACGAGAGCAAGAGAG 3339
Qy 2200 ATCCGTCTCTCTCTTTTAAACCTGGATCATTTCCCTGTGTAAAGCATTTTGAAGAC 2259
Db 3340 ATCCGTCTCTCTCTTTTAAACCTGGATCATTTCCCTGTGTAAAGCATTTTGAAGAC 3399
Qy 2260 AAGTACATATCTTTTCAAGCAAGTGGCAAGTTCACAGATTTTGTGACACAGAG 2319
Db 3400 AAGTACATATCTTTTCAAGCAAGTGGCAAGTTCACAGATTTTGTGACACAGAG 3459
Qy 2320 CTGGGCTCTCTCTGATGATTCATCCAAATTCACAGAGTGGTGAATGTCATCC 2379
Db 3460 CTGGGCTCTCTCTGATGATTCATCCAAATTCACAGAGTGGTGAATGTCATCC 3519
Qy 2380 TTTGGGGGAGTAACTTGAAGC 2401
Db 3520 TTTGGGGGAGTAACTTGAAGC 3541

RESULT 6

US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 93.7%; Score 2250; DB 13; Length 4966;
Best Local Similarity 94.5%; Pred. No. 0; Mismatches 0; Indels 141; Gaps 1;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

Qy 1 GGCAGTTCATTGATGAGAGTGAAGTAACTGACCGTTATCAACAGCTTTAGAGAA 60
Db 1757 GGCAGTTCATTGATGAGAGTGAAGTAACTGACCGTTATCAACAGCTTTAGAGAA 1816

Qy 61 GTATATCTGTGCTCTTCTTCTGCTGAGAGACATTTGCAAGCAAGAGAGATTTCTAAT 120
Db 1817 GTATATCTGTGCTCTTCTTCTGCTGAGAGACATTTGCAAGCAAGAGAGATTTCTAAT 1876
Qy 121 GATGTGAGAGGTGTGAAGAGACAGTTTCACTCATGAGGGGTACATGATGATTTGACA 180
Db 1877 GATGTGAGAGGTGTGAAGAGACAGTTTCACTCATGAGGGGTACATGATGATTTGACA 1936
Qy 181 GCCCATCAGAGCCGGGTGTGTATATTTCTAATTTGAGAGTGAATTTGAAACAGCA 240
Db 1937 GCCCATCAGAGCCGGGTGTGTATATTTCTAATTTGAGAGTGAATTTGAAACAGCA 1996
Qy 241 AATTTATCAAGATGAAGAACTGAAGTGAAGAGAGAGATGATCTCTTAATTCAGA 300
Db 1997 AATTTATCAAGATGAAGAACTGAAGTGAAGAGAGATGATCTCTTAATTCAGA 2056
Qy 301 TGGGAATGCTCAGAGGTGATGATGATGAAACAAAGCAATTTACATAGATTTTAATG 360
Db 2057 TGGGAATGCTCAGAGGTGATGATGATGAAACAAAGCAATTTACATAGATTTTAATG 2116
Qy 361 GATTTCAAGATGAGAACTGAAGATTTGAATGATCTGCTTACAAACAAAGAGAGAA 420
Db 2117 GATTTCAAGATGAGAACTGAAGATTTGAATGATCTGCTTACAAACAAAGAGAGAA 2176
Qy 421 ACAAGGAAATGAGAGAGAGCTCTGGAACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
Db 2177 ACAAGGAAATGAGAGAGAGCTCTGGAACCTGATCTTGAAGACCTTAAACGCCAAGTA 2236
Qy 481 CAACAACTATAGGTGTTCAAGAGATCTAGAACAAAGCAAGTCAAGGTCAATCTCTC 540
Db 2237 CAACAACTATAGGTGTTCAAGAGATCTAGAACAAAGCAAGTCAAGGTCAATCTCTC 2296
Qy 541 ACTCATAGTGTGTGTATGATGATCTGATGATGATGATGATGATGATGATGATGATG 600
Db 2297 ACTCATAGTGTGTGTATGATGATCTGATGATGATGATGATGATGATGATGATGATG 2356
Qy 601 GAACAATTAAGTATTTGGAGATGATGAGGCAACATCTGTAGATGAACAGAGACCGC 660
Db 2357 GAACAATTAAGTATTTGGAGATGATGAGGCAACATCTGTAGATGAACAGAGACCGC 2416
Qy 661 TGGGTTCTTTTCAAGAC----- 678
Db 2417 TGGGTTCTTTTCAAGACCGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 2476
Qy 679 ----- 678
Db 2477 CCTACTCAAGCTGTTACTGTGTGACACAACTGTGTGTTACTTAAGAAACTGCCATCTCC 2536
Qy 679 -----ACTCATAGATTTACTGCAAG 699
Db 2537 AAACTAGAATGCAATCTTCTTGTATGTTGAGTACCTGATCTCATAGATTTACTGCAAG 2596
Qy 700 TTCCCGCTGAGCTGGAAGATTTCTGTGCTGCTGCTTACAGAGTGAACCACTGCCAT 759
Db 2597 TTCCCGCTGAGCTGGAAGATTTCTGTGCTGCTGCTTACAGAGTGAACCACTGCCAT 2656
Qy 760 GTCTTACAGATGCTACCCGTGAAGAGAGCTCTTGAAGAGCTTCAAGGAGTAAAGAG 819
Db 2657 GTCTTACAGATGCTACCCGTGAAGAGAGCTCTTGAAGAGCTTCAAGGAGTAAAGAG 2716
Qy 820 CTGATTAACATGCTGAAGACCTTCAAGGTGAATTTGAAGCTCAAGAGATTTATCAG 879
Db 2717 CTGATTAACATGCTGAAGACCTTCAAGGTGAATTTGAAGCTCAAGAGATTTATCAG 2776
Qy 880 AACCTGATGAAGAAACAGCAAAATCTGAGATCTGGAAGTTCGATGATGACAGTC 939
Db 2777 AACCTGATGAAGAAACAGCAAAATCTGAGATCTGGAAGTTCGATGATGACAGTC 2836
Qy 940 CTGTTACAAAGAGCTTTGATTAATGAATTTCAAGGTGAAGTGAAGTTCGGAAGAGTCT 999
Db 2837 CTGTTACAAAGAGCTTTGATTAATGAATTTCAAGGTGAAGTGAAGTTCGGAAGAGTCT 2896
Qy 1000 CTCACATTAAGTCCCATTTGGAAGCAAGTCTGACAGTGAAGACGCTGACACTTCT 1059

Db 2897 CTCACATTAAGGCGCATTTGGAAGCCAGTTCTGACCGATGGAAGGCTGTGACCTTCT 2956
Qy 1060 CTGACAGAACTTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1119
Db 2957 CTGACAGAACTTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3016
Qy 1120 ATTGAGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179
Db 3017 ATTGAGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3076
Qy 1180 GAATGAAAACCTTAAGACCTGTAATCATGAGTACTTGTGACCTGTAATGTAATTTCTG 1239
Db 3077 GAATGAAAACCTTAAGACCTGTAATCATGAGTACTTGTGACCTGTAATGTAATTTCTG 1316
Qy 1240 ACAGAGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1299
Db 3137 ACAGAGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3196
Qy 1300 GAGAGAGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1359
Db 3197 GAGAGAGAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3256
Qy 1360 GAGTGGGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1419
Db 3257 GAGTGGGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3316
Qy 1420 GAAAAGCTCCAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1479
Db 3317 GAAAAGCTCCAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3376
Qy 1480 GAGTGGATCAAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1539
Db 3377 GAGTGGATCAAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3436
Qy 1540 CACTTCGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1599
Db 3437 CACTTCGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3496
Qy 1600 CAGCTCAATGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1659
Db 3497 CAGCTCAATGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3556
Qy 1660 CTGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1719
Db 3557 CTGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3616
Qy 1720 CGAGTCAAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1779
Db 3617 CGAGTCAAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3676
Qy 1780 TCAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1839
Db 3677 TCAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3736
Qy 1840 ATCAAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1899
Db 3737 ATCAAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3796
Qy 1900 TCTTTAGCTGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1959
Db 3797 TCTTTAGCTGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3856
Qy 1960 AGACTGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2019
Db 3857 AGACTGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3916
Qy 2020 GACGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2079
Db 3917 GACGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3976
Qy 2080 TTGAGAGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2139

Db 3977 TTGACCACTATTATATGACCGCTGAGCAAGACCAACATTTGCTCAAGTCCCTTC 4036
Qy 2140 TGCGTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2199
Db 4037 TGCGTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4096
Qy 2200 ATCCGTCTCTGCTTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2259
Db 4097 ATCCGTCTCTGCTTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4156
Qy 2260 AAGTACAGATACCTTTTCAAGAGTGGCAAGTTCAACAGATTTTGTGACAGCGCAG 2319
Db 4157 AAGTACAGATACCTTTTCAAGAGTGGCAAGTTCAACAGATTTTGTGACAGCGCAG 4216
Qy 2320 CTGGGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2379
Db 4217 CTGGGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4276
Qy 2380 TTTGGGGAGTAACTTGAAC 2401
Db 4277 TTTGGGGAGTAACTTGAAC 4298

RESULT 7
US-09-845-416-34
Sequence 34, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DEL142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 34
LENGTH: 4990
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-34

Query Match 93.7%; Score 2250; DB 13; Length 4990;
Best Local Similarity 94.5%; Pred. No. 0; Indels 141; Gaps 1;
Matches 2401; Conservative 0; Mismatches 0

Qy 1 GGCAGTTCATTGATGAGAGTGAAGTAACTGACCGTTATCAACAGCTTTAGAGAA 60
Db 1781 GGCAGTTCATTGATGAGAGTGAAGTAACTGACCGTTATCAACAGCTTTAGAGAA 1840
Qy 61 GTATTATCGTGTCTTCTTCTGCTGAGACATTTGCAAGCAAGAGAGATTTTAT 120
Db 1841 GTATTATCGTGTCTTCTTCTGCTGAGACATTTGCAAGCAAGAGAGATTTTAT 1900
Qy 121 GATGGAAGTGTGTAAGACAGCTTCACTCATGAGGGGTATCATGATTTTGA 180
Db 1901 GATGGAAGTGTGTAAGACAGCTTCACTCATGAGGGGTATCATGATTTTGA 1960
Qy 181 GCCCATCAGGGCCGGTGGTAATTTCAATTTGGGAAGTGAAGTGAACAG 240
Db 1961 GCCCATCAGGGCCGGTGGTAATTTCAATTTGGGAAGTGAAGTGAACAG 2020
Qy 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAAGACAGATGATCTCTTAATTCA 300
Db 2021 AAATTATCAGAAGATGAAGAACTGAAGTACAAAGACAGATGATCTCTTAATTCA 2080
Qy 301 TGGGAATGCTCAGGGTACCTAGCATGAAAAAACAAGCAATTTCATGAGTTTAATG 360
Db 2081 TGGGAATGCTCAGGGTACCTAGCATGAAAAAACAAGCAATTTCATGAGTTTAATG 2140
Qy 361 GATCTCAAGATATGAAGAACTGAAAGGTTGAATGACTGCTAAACAAAACAGAGAA 420

[illegible]

322	GAGGAGAGAGCCCAAGATATGCATCTGGCTTCTACGAAAGAGAGCTGAGAGTCAATACT	3280
QY	1360 GAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGAGAGAAAAATAGATGAGACCTT	1419
Db	3281 GAGTGGGAAAAATTGAACCTGCACCTCCGCTGACTGGAGAGAAAAATAGATGAGACCTT	3340
QY	1420 GAAAGACTCCAGGAACTTCAAGAGGCCAGGATGAGCTGACCTCAAGCTGGCCAACT	1479
Db	3341 GAAAGACTCCAGGAACTTCAAGAGGCCAGGATGAGCTGACCTCAAGCTGGCCAACT	3400
QY	1480 GAGGTATTCAGGGATCTTGGCAGCCCGTGGGGAATCTCTCATTTGATCTCTTCCAGAT	1539
Db	3401 GAGGTATTCAGGGATCTTGGCAGCCCGTGGGGAATCTCTCATTTGATCTCTTCCAGAT	3460
QY	1540 CACCTCGAAGAAAGTCAGGACACTTCGAGAGAAATTCGCGCTCGAAGAGAAAGTATGAC	1599
Db	3461 CACCTCGAAGAAAGTCAGGACACTTCGAGAGAAATTCGCGCTCGAAGAGAAAGTATGAC	3520
QY	1600 CAGCTCAATGACTTGTCTGCGCAGCTTACCACTTTGGGCAATTCAGCTTCAACCGTATAC	1659
Db	3521 CAGCTCAATGACTTGTCTGCGCAGCTTACCACTTTGGGCAATTCAGCTTCAACCGTATAC	3580
QY	1660 CTCAGCACTCTGGAAGACCTGGAACACAGATGGAGGTTCTGCAAGTGGCGGTGAGAGAC	1719
Db	3581 CTCAGCACTCTGGAAGACCTGGAACACAGATGGAGGTTCTGCAAGTGGCGGTGAGAGAC	3640
QY	1720 CGAGTCAAGCAGCTGCAATGAAGCCACAGAGGACTTTTGGTCCAGCATCTCAGACTTCTT	1779
Db	3641 CGAGTCAAGCAGCTGCAATGAAGCCACAGAGGACTTTTGGTCCAGCATCTCAGACTTCTT	3700
QY	1780 TTCACGCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGGCCAACAAGAGCCCTACAT	1839
Db	3701 TTCACGCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGGCCAACAAGAGCCCTACAT	3760
QY	1840 ATCAACACGAGACTCTCAACCAACTTGTGGGACCATCCAAAATGACAGAGCTTACAG	1899
Db	3761 ATCAACACGAGACTCTCAACCAACTTGTGGGACCATCCAAAATGACAGAGCTTACAG	3820
QY	1900 TCTTTAGCTGACTGAAATATGTCAATTTCACTTAATAGAGACTGCCATGAACCTCGA	1959
Db	3821 TCTTTAGCTGACTGAAATATGTCAATTTCACTTAATAGAGACTGCCATGAACCTCGA	3880
QY	1960 AGACTGAGAGAGGCCCTTGGTGGATCTCTGAGCTGTGACGTGCATGATGATGCTTG	2019
Db	3881 AGACTGAGAGAGGCCCTTGGTGGATCTCTGAGCTGTGACGTGCATGATGATGCTTG	3940
QY	2020 GACCAGCAACAACCTCAAGCAAAATGACACAGCCATGAGATCTCGCAATTAATATGT	2079
Db	3941 GACCAGCAACAACCTCAAGCAAAATGACACAGCCATGAGATCTCGCAATTAATATGT	4000
QY	2080 TTGACCACTATTATGACCGCTGGAGACAGAGCAAAATTGGTCAACGTCCCTCTC	2139
Db	4001 TTGACCACTATTATGACCGCTGGAGAGAGAGCAAAATTGGTCAACGTCCCTCTC	4060
QY	2140 TGGCTGATATGATCTGAACTGGCTCTGATGTTATGATAGAGGACGAACAGGAGG	2199
Db	4061 TGGCTGATATGATGATCTGAACTGGCTCTGATGTTATGATAGAGGACGAACAGGAGG	4120
QY	2200 ATCCGTCTCTGCTTTTAAAACCTGGCATCTTCCCTGTGTAAAGCAATTTGGAGAGAC	2259
Db	4121 ATCCGTCTCTGCTTTTAAAACCTGGCATCTTCCCTGTGTAAAGCAATTTGGAGAGAC	4180
QY	2260 AAGTACAGATATCTTTCAAGCAGTGGCAAGTCAACAGGATTTGTGACACAGCGCAGG	2319
Db	4181 AAGTACAGATATCTTTCAAGCAGTGGCAAGTCAACAGGATTTGTGACACAGCGCAGG	4240
QY	2320 CTGGGCTCTCTTTCGATGATTCATCCAAATTCAGAGACGTTGGTGAAGTTGCACTC	2379
Db	4241 CTGGGCTCTCTTTCGATGATTCATCCAAATTCAGAGACGTTGGTGAAGTTGCACTC	4300
QY	2380 TTGAGGAGCAATACATGAGC 2401	
Db	4301 TTGAGGAGCAATACATGAGC 4322	

RESULT 8
 US-09-845-416-2
 ; Sequence 2, Application US/09845416
 ; Publication No. US20030171312A1
 ; GENERAL INFORMATION:
 ; APPLICANT: XIAO, XIAO
 ; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
 ; FILE REFERENCE: DEL142
 ; CURRENT APPLICATION NUMBER: US/09/845,416
 ; CURRENT FILING DATE: 2001-04-30
 ; PRIOR APPLICATION NUMBER: 60/200,777
 ; PRIOR FILING DATE: 2000-04-28
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 4182
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-845-416-2

Query Match 86.1%; Score 2067; DB 13; Length 4182;
 Best Local Similarity 88.1%; Pred. No. 0;
 Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

QY 1 GGCAATTCATTGATGAGAGTGAAGTAACTGACCGTTATCAAAACGCTTTGAGAGAA 60
 DB 1000 GGCAATTCATTGATGAGAGTGAAGTAACTGACCGTTATCAAAACGCTTTGAGAGAA 1059
 QY 61 GATATTCGTGGCTCTTCTGCTGAGACACATTCGACAGCAAGAGAGATTTCTAT 120
 DB 1060 GATATTCGTGGCTCTTCTGCTGAGACACATTCGACAGCAAGAGAGATTTCTAT 1119
 QY 121 GATGAGAGAGTGGTGAAGACCAATTCATCTCATGAGGGTACATGATGATTTGACA 180
 DB 1120 GATGAGAGAGTGGTGAAGACCAATTCATCTCATGAGGGTACATGATGATTTGACA 1179
 QY 181 GCCCATCAGGCGGGGTGGTAAATTTCTACAAATTGGGAGATGATGATGAGAGAGAA 240
 DB 1180 GCCCATCAGGCGGGGTGGTAAATTTCTACAAATTGGGAGATGATGATGAGAGAGAA 1239
 QY 241 AATATATCAGAGATGAGAGAACTGAAGTACAGAGCAGATGATCTCTTAATTCAGA 300
 DB 1240 AATATATCAGAGATGAGAGAACTGAAGTACAGAGCAGATGATCTCTTAATTCAGA 1239
 QY 301 TGGGANTGCTCAAGGTAGCTGCAATGAAAAACAAGCAATTTACATAGAGTTTAATG 360
 DB 1300 TGGGANTGCTCAAGGTAGCTGCAATGAAAAACAAGCAATTTACATAGAGTTTAATG 1359
 QY 361 GATCTCCAGATCAGAACTGAAAGTGAATGACTGAGCTAACAAACAGAGAGAGAA 420
 DB 1360 GATCTCCAGATCAGAACTGAAAGTGAATGACTGAGCTAACAAACAGAGAGAGAA 1419
 QY 421 ACGAGAAAAAGAGAGAGAGCTCTTGGACCTGATCTTGAAGACTTAAAAAGCCCAAGTA 480
 DB 1420 ACGAGAAAAAGAGAGAGAGCTCTTGGACCTGATCTTGAAGACTTAAAAAGCCCAAGTA 1479
 QY 481 CAACAACATAGAGTGGCTTCAAGAGATCTAGAACAGAGAAAGATGAGAGGTAAATTCCTC 540
 DB 1480 CAACAACATAGAGTGGCTTCAAGAGATCTAGAACAGAGAAAGATGAGAGGTAAATTCCTC 1539
 QY 541 ACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 1540 ACTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1539
 QY 601 GAAACAACCTTAAGGTATGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 DB 1600 GAAACAACCTTAAGGTATGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1659
 QY 661 TGGGTTCTTTTACAGAGC----- 678

DB 1660 TGGGTTCTTTTACAGAGATCTTCTCAAAATGAGAAAGCTTTACTGAAAGACAGTGCCTT 1719
 QY 679 ----- 678
 DB 1720 TTTAGTGCATGCTTTTCAAAAAAGAGATGAGTGAACAGATTCACACACTGGCTTT 1779
 QY 679 ----- 678
 DB 1780 AAGATCAAAATGAAATGTTATCAAGTCTTCAAAAATGCGCTTTTAAAGCGATCTA 1839
 QY 679 ----- 678
 DB 1840 GAAAGAAAAAGCAATCCATGCGCAAACTGTATTCATCAACAAGATCTTCTTTCACA 1899
 QY 679 ----- 678
 DB 1900 CTGAGAAATPAGTCAGTACCCAGAGAGAGAGAGATGCTGATTAATTTGCCGCTGT 1959
 QY 679 -----ACTGATGATTAAGTCA 696
 DB 1960 TGGGATTAATTTAGTCCAAAACTTGAAGAGAGTACAGACAGACTCATGATTAATTCGAA 2019
 QY 697 CAGTTCCCTGAGACCTGGAAGAGTTCTTGCTGCTTACAGAGCTGAAACACTGCC 756
 DB 2020 CAGTTCCCTGAGACCTGGAAGAGTTCTTGCTGCTTACAGAGCTGAAACACTGCC 2079
 QY 757 AATGCTCTACAGATGCTACCCCTTAAGAGAAAGCTCTTGAAGACTCTCCAAAGAGTAAAA 816
 DB 2080 AATGCTCTACAGATGCTACCCCTTAAGAGAAAGCTCTTGAAGACTCTCCAAAGAGTAAAA 2139
 QY 817 GAGCTGATGAACAAATGAGAGAGCTCAAGGTGAATGAATGAAGTCAACAGATGTTAT 876
 DB 2140 GAGCTGATGAACAAATGAGAGAGCTCAAGGTGAATGAATGAAGTCAACAGATGTTAT 2139
 QY 877 CACACCTGATGATAAAAAGCCAAAAATCTGAGATCCCTGAGAGCTTCGATGATGCA 936
 DB 2200 CACACCTGATGATAAAAAGCCAAAAATCTGAGATCCCTGAGAGCTTCGATGATGCA 2259
 QY 937 GTCCTGTTACAAAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 996
 DB 2260 GTCCTGTTACAAAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2319
 QY 997 TCTCTCAACATTAGTCCCATTTGGAAGCAAGTCTGACAGTGAAGAGCTTCACCTT 1056
 DB 2320 TCTCTCAACATTAGTCCCATTTGGAAGCAAGTCTGACAGTGAAGAGAGCTTCACCTT 2379
 QY 1057 TCTCTCAAGAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
 DB 2380 TCTCTCAAGAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2439
 QY 1117 CTAATTGAGAGCACTTCCAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1176
 DB 2440 CTAATTGAGAGCACTTCCAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2499
 QY 1177 AGGGAATGAAAATTAAGAACTGTAATCATGATGATGATGATGATGATGATGATGATGAT 1236
 DB 2500 AGGGAATGAAAATTAAGAACTGTAATCATGATGATGATGATGATGATGATGATGATGAT 2559
 QY 1237 CTGACAGAGAGCTTCTGAGAGAGCTGAGAGAACTTACACAGAGAGAGAGAGAGAGAGAG 1296
 DB 2560 CTGACAGAGAGCTTCTGAGAGAGCTGAGAGAACTTACACAGAGAGAGAGAGAGAGAGAG 2619
 QY 1297 CTTGAG 1356
 DB 2620 CTTGAG 2679
 QY 1357 ACTGAGTGGGAAAAATTAAGAACTGCACTGCGCTGAGCTGAGAGAGAGAGAGAGAGAGAG 1416
 DB 2680 ACTGAGTGGGAAAAATTAAGAACTGCACTGCGCTGAGCTGAGAGAGAGAGAGAGAGAGAG 2739
 QY 1417 CTTGAG 1476
 DB 2740 CTTGAG 2799

QY 1477 GCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATGTACTCTCCAA 1536
DB GCTGAGTGATCAAGGGATCCTGGCAGCCCGTGGCGATCTCTCATGTACTCTCCAA 2859
QY 1537 GATCACTCCGAGAAAGTCAAGGCACTTCGAGAGAAATTTGGGCTCTGAAAGAAAGCTG 1596
DB GATCACTCCGAGAAAGTCAAGGCACTTCGAGAGAAATTTGGGCTCTGAAAGAAAGCTG 2919
QY 1597 AGCCAGTCAATGACCTTGTCTGCGACCTTACCACTTTGGGCACTTCTCACCCTAT 1656
DB AGCCAGTCAATGACCTTGTCTGCGACCTTACCACTTTGGGCACTTCTCACCCTAT 2979
QY 1657 AACCTCAGCACTTCGAGAAAGCTTGAACCAAGAGTGAAGCTTTCGAGAGTGGCTGAG 1716
DB AACCTCAGCACTTCGAGAAAGCTTGAACCAAGAGTGAAGCTTTCGAGAGTGGCTGAG 3039
QY 1717 GACCGAGTCAGGCACTGCATGAGAGCCCAAGGCACTTTGGTCCAGATCTCAGCACTTT 1776
DB GACCGAGTCAGGCACTGCATGAGAGCCCAAGGCACTTTGGTCCAGATCTCAGCACTTT 3099
QY 1777 CTTTCCAGCTCTGTCCAGGGTCTCTGAGAGAGCCATCTTGCCTAAACAAAGTCCCTAC 1836
DB CTTTCCAGCTCTGTCCAGGGTCTCTGAGAGAGCCATCTTGCCTAAACAAAGTCCCTAC 3159
QY 1837 TATATCAACAGAGCACTCAACCAACTTGTGGGACCATCCCAAAATGACAGAGCTTAC 1896
DB TATATCAACAGAGCACTCAACCAACTTGTGGGACCATCCCAAAATGACAGAGCTTAC 3219
QY 1897 CAGCTTTAGTCACTGATATATGTCAATTTCTAGCTTATAGAGCTGCCATGAACTC 1956
DB CAGCTTTAGTCACTGATATATGTCAATTTCTAGCTTATAGAGCTGCCATGAACTC 3279
QY 1957 CGAAGCTGCAAGAGCCCTTTGCTGATCTCTTGAAGCTCTGAGCTGCACTGATGCC 2016
DB CGAAGCTGCAAGAGCCCTTTGCTGATCTCTTGAAGCTCTGAGCTGCACTGATGCC 3339
QY 2017 TTGAGACAGCAACCTCAAGCAAAATGACAGCCCATGATATCTGCAAGATTTAT 2076
DB TTGAGACAGCAACCTCAAGCAAAATGACAGCCCATGATATCTGCAAGATTTAT 3399
QY 2077 TGTGTTGACCACTTTATGACCGCTGAGCAAGCAACCAATTGCTGATGCTCCT 2136
DB TGTGTTGACCACTTTATGACCGCTGAGCAAGCAACCAATTGCTGATGCTCCT 3459
QY 2137 CTCTGGTGGATATGATGTCTGACTGAGCTGCTGATGTTTATGATAGCGGAGCAAG 2196
DB CTCTGGTGGATATGATGTCTGACTGAGCTGCTGATGTTTATGATAGCGGAGCAAG 3519
QY 2197 AGGATCCGTCCTGTCTTTTAAACCTGECATCATTTCCCTGTGAAAGCAATTGGAA 2256
DB AGGATCCGTCCTGTCTTTTAAACCTGECATCATTTCCCTGTGAAAGCAATTGGAA 3579
QY 2257 GACAGTACAGATACCTTTTCAAGCAAGGCAAGTCAAGAGATTTTGGACAGAGC 2316
DB GACAGTACAGATACCTTTTCAAGCAAGGCAAGTCAAGAGATTTTGGACAGAGC 3639
QY 2317 AGGCTGGGCTCTCTGATGATTTCAACAAATTCACAGACAGTGGGATGAACTTGA 2376
DB AGGCTGGGCTCTCTGATGATTTCAACAAATTCACAGACAGTGGGATGAACTTGA 3699
QY 2377 TCTTTTGGGGGAGTAACTTGAGC 2401
DB TCTTTTGGGGGAGTAACTTGAGC 3724

RESULT 9

US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

;; TITLE OF INVENTION: THEREOF
;; FILE REFERENCE: DE1142
;; CURRENT APPLICATION NUMBER: US/09/845,416
;; CURRENT FILING DATE: 2001-04-30
;; PRIOR APPLICATION NUMBER: 60/200,777
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 27
;; LENGTH: 5149
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-845-416-27
Query Match 86.1%; Score 2067; DB 13; Length 5149;
Best Local Similarly 86.1%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;
QY 1 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAA 60
DB 1757 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAA 1816
QY 61 GATATATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCAAGAGAGATTCTAT 120
DB 1817 GATATATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCAAGAGAGATTCTAT 1876
QY 121 GATGCGAAGTGTGAAAGCCAGTTTACTCATGAGGGGTACATGATGATTTGACA 180
DB 1877 GATGCGAAGTGTGAAAGCCAGTTTACTCATGAGGGGTACATGATGATTTGACA 1936
QY 181 GCCCATCAGGGCGGGTGGTATATTTCTACATTTGGAGTGAAGTATGGAACAGGA 240
DB 1937 GCCCATCAGGGCGGGTGGTATATTTCTACATTTGGAGTGAAGTATGGAACAGGA 1996
QY 241 AATATTCAGAAATGAGAAACCTGAATGACAGAGCAAGTATCTCTAAATTCAGA 300
DB 1997 AATATTCAGAAATGAGAAACCTGAATGACAGAGCAAGTATCTCTAAATTCAGA 2056
QY 301 TGGGAATGCTCAGGGTACTGATGATGAAACCAACCAATTATCATAGATTATG 360
DB 2057 TGGGAATGCTCAGGGTACTGATGATGAAACCAACCAATTATCATAGATTATG 2116
QY 361 GATCTCAGAAATCAGAACTGAAGAGTTGAATGACTGCTTAACAAAAACAGAAAGAA 420
DB 2117 GATCTCAGAAATCAGAACTGAAGAGTTGAATGACTGCTTAACAAAAACAGAAAGAA 2176
QY 421 ACAAGGAAATGAGAGAGAGCTCTGGAAGCTGATCTTGAAGCTTAAACGCCAAGTA 480
DB 2177 ACAAGGAAATGAGAGAGAGCTCTGGAAGCTGATCTTGAAGCTTAAACGCCAAGTA 2236
QY 481 CAACCAATAGGTGCTTCAAGAAATCTAGAACAGAACAAAGTCAAGGTCATTTCTTC 540
DB 2237 CAACCAATAGGTGCTTCAAGAAATCTAGAACAGAACAAAGTCAAGGTCATTTCTTC 2296
QY 541 ACTCAGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 2297 ACTCAGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2356
QY 601 GAACCACTTAAAGTATGAGAGATGATGAGGCAACATCTGATGATGATGATGATGATGAT 660
DB 2357 GAACCACTTAAAGTATGAGAGATGATGAGGCAACATCTGATGATGATGATGATGATGAT 2416
QY 661 TGGGTTCTTTTACAGAC----- 678
DB 2417 TGGGTTCTTTTACAGACCTCTTCAATAGCAACGTTTATCTGAGAGACAGTGCCTT 2476
QY 679 ----- 678
DB 2477 TTTAGTCAATGCTTTCAAGAAAGAGATGAGAGATTCACACATCTGCTTT 2536
QY 679 ----- 678
DB 2537 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACTGGCGTTTAAAGCGGATCTA 2596

SOFTWARE: PatentIn version 3.1
 SEQ ID NO 41
 LENGTH: 5462
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthetic
 US-10-149-736-41

Query Match 76.6%; Score 1839.4; DB 13; Length 5462;
 Best Local Similarity 88.1%; Pred. No. 0;
 Matches 2115; Conservative 0; Mismatches 76; Indels 210; Gaps 2;

QY 1 GGCAGTTCATTGATGAGAGTGAAGTAACTTGACCGTTTCAACAGCTTTGAGAA 60
 DB 1199 GGCAGTTCATTGATGAGAGTGAAGTAACTTGACCGTTTCAACAGCTTTGAGAA 1258
 QY 61 GTATTATCGTGGCTCTTCTGCTGAGGACATTTGACAGCAAGAGAGATTCTAT 120
 DB 1259 GTATTATCGTGGCTCTTCTGCTGAGGACATTTGACAGCAAGAGAGATTCTAT 1318
 QY 121 GATGTGAAGTGTGTGAAGACCAAGTTTCACTCATGAGGGTACATGATGATTTGACA 180
 DB 1319 GATGTGAAGTGTGTGAAGACCAAGTTTCACTCATGAGGGTACATGATGATTTGACA 1378
 QY 181 GCCCATCAGGGCCGGTGTGTATATCTTACATTTGGAGATGATGATTTGAGAACGGA 240
 DB 1379 GCCCATCAGGGCCGGTGTGTATATCTTACATTTGGAGATGATGATTTGAGAACGGA 1438
 QY 241 AATTATCAGAGATGAGAGAACTGAGTACAGAGAGATGATGATTTCTTAATTCAAGA 300
 DB 1439 AATTATCAGAGATGAGAGAACTGAGTACAGAGAGATGATGATTTCTTAATTCAAGA 1498
 QY 301 TGGGAATGCTTCAAGGTAGCTTACATGAGAAAAACAAGCAATTTCATAGCTTTTATG 360
 DB 1499 TGGGAATGCTTCAAGGTAGCTTACATGAGAAAAACAAGCAATTTCATAGCTTTTATG 1558
 QY 361 GATCTCCAGTAATCAGAACTGAAGATTGATGATGCTGGTCAACAAAAAGAGAAAGA 420
 DB 1559 ACCA----- 1562
 QY 421 ACAAGGAAATGAGAGAGAGCGCTTTGACCTGATCTTGAAGACCTTAAAGCCCAAGTA 480
 DB 1563 -----CTATTGGAGCGCTCTCTACTCAAGACTGTTACTCTGCTG 1600
 QY 481 CAACAACATTAAGTGTTCAGAGAATCTAGAACAAAGAAAGTCAAGTCAATTCTCTC 540
 DB 1601 ACAACAACCTGTGTTTCTAGAGAACTGCCATCTCCAACTTAGAATGCCATCTTCTTG 1660
 QY 541 ACTCAATGTGTGTAGTGTAGTATCTAGTGAAGATCAAGCAACTGCTCTTGGAA 600
 DB 1661 AGTTGGAG----- 1669
 QY 601 GAACAACCTTAAGGATTGGGAGATGATGGGCAACATCTGTAGATGACAGAGAACCCG 660
 DB 1670 ----- 1669
 QY 661 TGGGTTCTTTTACAAGACACTATAGATTACTGACAAGTCCCGCTGACCTGAGAAAG 720
 DB 1670 -----CATAGATTACTGCAACAGTTCCCGCTGAGCTGAGAAAG 1708
 QY 721 TTTCTTGCTGTGCTTACAGAGCTGAAACAATGCCAATGCTTACAGAGTCTACCCGT 780
 DB 1709 TTTCTTGCTGTGCTTACAGAGCTGAAACAATGCCAATGCTTACAGAGTCTACCCGT 1768
 QY 781 AAGGAAGGCTCTTAAGACTCAAGAGGAGTAAAGAGCTGATGAAACAATGGCAAGC 840
 DB 1769 AAGGAAGGCTCTTAAGACTCAAGAGGAGTAAAGAGCTGATGAAACAATGGCAAGC 1828
 QY 841 CTCCAAGTGAATTGAGCTCAACAAGATGTTTATCACAACCTGATGAAACAAGCCAA 900
 DB 1829 CTCCAAGTGAATTGAGCTCAACAAGATGTTTATCACAACCTGATGAAACAAGCCAA 1888

QY 901 AAATCCTGAGATCCCTGAGAGTTCCTGATGATGAGTCCCTTCAAAAGACCTTGAT 960
 DB 1889 AAATCCTGAGATCCCTGAGAGTTCCTGATGATGAGTCCCTTCAAAAGACCTTGAT 1948
 QY 961 AACATGACTTCAAGTGAAGTGAACCTTGGAAAAAGTCTCTCAACATTAGTCCCATTTG 1020
 DB 1949 AACATGACTTCAAGTGAAGTGAACCTTGGAAAAAGTCTCTCAACATTAGTCCCATTTG 2008
 QY 1021 GAAGCAGTTCCTGACAGTGAAGGGTCTGACCTTTCTGACAGAACTTGTGTGG 1080
 DB 2009 GAAGCAGTTCCTGACAGTGAAGGGTCTGACCTTTCTGACAGAACTTGTGTGG 2068
 QY 1081 CTACAGCTGAAGATGATGAATTAAAGCCGAGGAGACCATTTAGAGGACCTTTCAGCA 1140
 DB 2069 CTACAGCTGAAGATGATGAATTAAAGCCGAGGAGACCATTTAGAGGACCTTTCAGCA 2128
 QY 1141 GTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGAATTGAAACTTAAGACCT 1200
 DB 2129 GTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGAATTGAAACTTAAGACCT 2188
 QY 1201 GTATATGATGATCTTGAAGTGTGAGATGATGATGATGATGATGATGATGATGATGAT 1260
 DB 2189 GTATATGATGATCTTGAAGTGTGAGATGATGATGATGATGATGATGATGATGATGAT 2248
 QY 1261 CTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTGAGAGAGAGCCAGATGTC 1320
 DB 2249 CTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTGAGAGAGAGCCAGATGTC 2308
 QY 1321 ACTGGCTTCTACGAAAGCAAGCTGAGAGAGTCAATCTGATGGGAAAAATTGAACCTG 1380
 DB 2309 ACTGGCTTCTACGAAAGCAAGCTGAGAGAGTCAATCTGATGGGAAAAATTGAACCTG 2368
 QY 1381 CACTCGCTGACTGCGAGAGAAAAATGATGAGACCTTGAAGAATCCAGGAACTTCA 1440
 DB 2369 CACTCGCTGACTGCGAGAGAAAAATGATGAGACCTTGAAGAATCCAGGAACTTCA 2428
 QY 1441 GAGGCCACGATGAGTGAACCTCAAGCTGCGCAAGCTGAGATGATCAAGGAACTTGG 1500
 DB 2429 GAGGCCACGATGAGTGAACCTCAAGCTGCGCAAGCTGAGATGATCAAGGAACTTGG 2488
 QY 1501 CAGCCCTGGGCAATCTCTCATGATCTCTCCAAAGATCACTGAGAAATCAAGGCA 1560
 DB 2489 CAGCCCTGGGCAATCTCTCATGATCTCTCCAAAGATCACTGAGAAATCAAGGCA 2548
 QY 1561 CTTCGAGAGAAATTCGCTCTGAAAGAAAGTGAAGCAAGTCAATGACCTTGTCTGC 1620
 DB 2549 CTTCGAGAGAAATTCGCTCTGAAAGAAAGTGAAGCAAGTCAATGACCTTGTCTGC 2608
 QY 1621 CAGCTTACCACTTGGGCAATGAGCTCTCAAGCTTACCTGAGCACTTGGAAAGCCTG 1680
 DB 2609 CAGCTTACCACTTGGGCAATGAGCTCTCAAGCTTACCTGAGCACTTGGAAAGCCTG 2668
 QY 1681 AACACCAATGAGAAAGCTTCTGAGAGTGGCGCTGAGAGACGAGTCAAGCTGATGA 1740
 DB 2669 AACACCAATGAGAAAGCTTCTGAGAGTGGCGCTGAGAGACGAGTCAAGCTGATGA 2728
 QY 1741 GCCCAGAGGACCTTGGTCCAGCATCTCAGCACTTCTTCAACGTCGTCCAGGATCCC 1800
 DB 2729 GCCCAGAGGACCTTGGTCCAGCATCTCAGCACTTCTTCAACGTCGTCCAGGATCCC 2788
 QY 1801 TGGGAGAGGCACTCGGCAAAAGAGGCTCTATATCAACAGAGACTCAAAACA 1860
 DB 2789 TGGGAGAGGCACTCGGCAAAAGAGGCTCTATATCAACAGAGACTCAAAACA 2848
 QY 1861 ACTTGTGGAGCCATCCAAATGACAGAGCTTACAGATCTTATAGCTGATGAATAT 1920
 DB 2849 ACTTGTGGAGCCATCCAAATGACAGAGCTTACAGATCTTATAGCTGATGAATAT 2908
 QY 1921 GTAGATTCAGACTTATAGACTGCCATGAAATCCGAAAGCTGAGAGAGGCTTTCG 1980
 DB 2909 GTAGATTCAGACTTATAGACTGCCATGAAATCCGAAAGCTGAGAGAGGCTTTCG 2968
 QY 1981 TTGAGTCTTGTGAGCTGTAGCTGATGATGCTTGGACAGCAACACTCAAGCA 2040

Db 2969 TTGATCTCTTGAGCTCTGACGCTGATGATGCTTGGACCAACCACTCAAGCA 3028
Qy 2041 AATGACCAAGCCATGGAATATCTCGAGATTATTTATTTATGACCACTATTTATGACCGC 2100
Db 3029 AATGACCAAGCCATGGAATATCTCGAGATTATTTATTTATGACCACTATTTATGACCGC 3088
Qy 2101 CTGGAGCAAGAGACAACTTTGTCACGTCCTCTGCTGGATGATGTCGTGTAAC 2160
Db 3089 CTGGAGCAAGAGACAACTTTGTCACGTCCTCTGCTGGATGATGTCGTGTAAC 3148
Qy 2161 TGGCTGCTGATTTTATGATACGAGACCAAGGAGGATCCGTCTCTTTTAA 2220
Db 3149 TGGCTGCTGATTTTATGATACGAGACCAAGGAGGATCCGTCTCTTTTAA 3208
Qy 2221 ACTGGATCATTTCCCTGTGTAAAGACATTGGAAAGACATGATACCTTTTCAAG 2280
Db 3209 ACTGGATCATTTCCCTGTGTAAAGACATTGGAAAGACATGATACCTTTTCAAG 3268
Qy 2281 CAAGTGGCAAGTTCAACAGGATTTTGTACACGAGGCTGGCTCTCTTCTGATGAT 2340
Db 3269 CAAGTGGCAAGTTCAACAGGATTTTGTACACGAGGCTGGCTCTCTTCTGATGAT 3328
Qy 2341 TCTATCCAAATTCCAAAGACATTTGGGTGAAGTTGCACTTTGGGGGCGATTAACATTGAG 2400
Db 3329 TCTATCCAAATTCCAAAGACATTTGGGTGAAGTTGCACTTTGGGGGCGATTAACATTGAG 3388
Qy 2401 C 2401
Db 3389 C 3389

RESULT 11

US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-10

Query Match 72.3%; Score 1737; DB 13; Length 3531;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 576 AGATCAGCGACTGCTCTTTGGAGAACAACTTAAGTATGGAGATCGATGGCAAA 635
Db 1248 AGAAGATGAAGAACTGAAGTACAAGACAGATGATCTCTAAATTCAAGATGGGAATG 1307
Qy 636 CATCTGATGATGAGCAAGAGACCGCTGCTTTTACAGACACTCATAGATTATCTGCA 695
Db 1308 CCTCAGGATGCTAGATGAAAAACAAGCAATTTACATAGATCTCATGATTACTGCA 1367
Qy 696 AAGTTCCCTGGACCTGAAAAAGTTTCTGCTGCTTACAGAGCTGAAAAACAATGC 755
Db 1368 AAGTTCCCTGGACCTGAAAAAGTTTCTGCTGCTTACAGAGCTGAAAAACAATGC 1427
Qy 756 CATGCTCTACAGAGATGATACCGCTAGAGAAAGCTCTTGAAGACTCCAGAGAGTAA 815
Db 1428 CATGCTCTACAGAGATGATACCGCTAGAGAAAGCTCTTGAAGACTCCAGAGAGTAA 1487

Qy 816 AGAGCTGATGAAACAATGGCAAGACCTCCAAAGTGAATTTGAAGTCTACACATGTTTA 875
Db 1488 AGAGCTGATGAAACAATGGCAAGACCTCCAAAGTGAATTTGAAGTCTACACATGTTTA 1547
Qy 876 TCAACAACCTGATGAAACAAGCCAAAAAATCTGAGATCCCTGGAAAGTTCCGATATGC 935
Db 1548 TCAACAACCTGATGAAACAAGCCAAAAAATCTGAGATCCCTGGAAAGTTCCGATATGC 1607
Qy 936 AGTCCCTGTACAAAAGACGTTTGAATACATGAATCTCAAGTGAATGAACTTGGAAAA 995
Db 1608 AGTCCCTGTACAAAAGACGTTTGAATACATGAATCTCAAGTGAATGAACTTGGAAAA 1667
Qy 996 GTCTCTCAACATTAGGCTCCATTGGAAAGCACTTGCAGATGGAAGGCTGCACT 1055
Db 1668 GTCTCTCAACATTAGGCTCCATTGGAAAGCACTTGCAGATGGAAGGCTGCACT 1127
Qy 1056 TTCTCTGACAGAACTTCTGCTGCTGCTACAGCTGAAAGATGATGAATTAAGCCGCAAGC 1115
Db 1128 TTCTCTGACAGAACTTCTGCTGCTGCTACAGCTGAAAGATGATGAATTAAGCCGCAAGC 1187
Qy 1116 ACCATTGGAGGCGACCTTTCAGAGCTTGAAGAGCAAGATGATCATAGAGGCTTTCA 1175
Db 1788 ACCATTGGAGGCGACCTTTCAGAGCTTGAAGAGCAAGATGATCATAGAGGCTTTCA 1847
Qy 1176 GAGGGAATTGAAAACTTAAGACCTGTATCATAGATCACTTGAATGTACGAATATT 1235
Db 1848 GAGGGAATTGAAAACTTAAGACCTGTATCATAGATCACTTGAATGTACGAATATT 1907
Qy 1236 TCTGACAGAGCAGCTTTTGAAGAGATGAGAAATCTTACAGAGCCCAAGAGACTGCC 1295
Db 1908 TCTGACAGAGCAGCTTTTGAAGAGATGAGAAATCTTACAGAGCCCAAGAGACTGCC 1967
Qy 1296 TCTGAGAGAGAGCCAGAAATGTCACTGAGCTTCTACGAAACAGAGCTGAGAGGTCA 1355
Db 1968 TCTGAGAGAGAGCCAGAAATGTCACTGAGCTTCTACGAAACAGAGCTGAGAGGTCA 2027
Qy 1356 TACTGAGTGGAAAAATTTGAACCTGCACTCCGCTGACTGCGACAGAAAAATAGATGAGC 1415
Db 2028 TACTGAGTGGAAAAATTTGAACCTGCACTCCGCTGACTGCGACAGAAAAATAGATGAGC 2087
Qy 1416 CTTTGAAGACCTCAGGAATTTCAAGAGGCAAGATGAGCTGAGACCTCAAGCTGGCCCA 1475
Db 2088 CTTTGAAGACCTCAGGAATTTCAAGAGGCAAGATGAGCTGAGACCTCAAGCTGGCCCA 2147
Qy 1476 AGCTGAGTATCAAGAGATCTGCGAGCCCGTGGCGATCTTCTCATTTGACTCTGTCCA 1535
Db 2148 AGCTGAGTATCAAGAGATCTGCGAGCCCGTGGCGATCTTCTCATTTGACTCTGTCCA 2207
Qy 1536 AGATCAGCTGAGAAAGTCAAGGCACTTGAAGAGAAATTTGGCCCTCTGAAAGAGACGT 1595
Db 2208 AGATCAGCTGAGAAAGTCAAGGCACTTGAAGAGAAATTTGGCCCTCTGAAAGAGACGT 2267
Qy 1596 GAGCAGAGTCAATGACCTTGTGCTGCGAGCTTACCACTTTGGGCACTCAGCTCACCGTA 1655
Db 2268 GAGCAGAGTCAATGACCTTGTGCTGCGAGCTTACCACTTTGGGCACTCAGCTCACCGTA 2327
Qy 1656 TAACTTCAAGCATCTTGGAAAGACTTGAACACAGATGAAAGCTTCTGAGAGTGGCGGTCA 1715
Db 2328 TAACTTCAAGCATCTTGGAAAGACTTGAACACAGATGAAAGCTTCTGAGAGTGGCGGTCA 2387
Qy 1716 GAGCAGAGTCAAGGAGTGAAGACCAAGGAGCTTTGGCCAGCATCTCAGACACT 1775
Db 2388 GAGCAGAGTCAAGGAGTGAAGACCAAGGAGCTTTGGCCAGCATCTCAGACACT 2447
Qy 1776 TCTTTTCAAGTGTCTCAGAGGCTCTTGGAGAGAGCATCTGCGCAACAAAGTGCCCTTA 1835
Db 2448 TCTTTTCAAGTGTCTCAGAGGCTCTTGGAGAGAGCATCTGCGCAACAAAGTGCCCTTA 2507
Qy 1836 CTATATCAACAGAGACCTCAACCACTTGTGGGACCATCCCAAAATGACAGAGCTTA 1895
Db 2508 CTATATCAACAGAGACCTCAACCACTTGTGGGACCATCCCAAAATGACAGAGCTTA 2567
Qy 1896 CCACTCTTTAGCTGACCTGAATATGATGATGATGATGATGATGATGATGATGATGAT 1955

Db	2568	CCAGTCTTTAGCTACCTCGAATTAATGTGAGATTTCTCAGCTTATATAGACCTGCGCAAGAACT	2627
Qy	1956	CCGAAAGCTGAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTACGTGATGTGATGC	2015
Db	2628	CCGAAAGACTGAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTACGTGATGTGATGC	2687
Qy	2016	CTTGACACGACACAACCTCAAGCAAAATGACACAGCCCATGATATCTCGACATTAATTA	2075
Db	2688	CTTGACACGACACAACCTCAAGCAAAATGACACAGCCCATGATATCTCGACATTAATTA	2747
Qy	2076	TTGTTTGAACCACTATTTATATGACCCGCTGACAGACAGACAACAATTTGTCACAGTCC	2135
Db	2748	TTGTTTGAACCACTATTTATATGACCCGCTGACAGACAGACAACAATTTGTCACAGTCC	2807
Qy	2136	TCTCTGCGTGGANATGTGTCTGAACTGGCTGCTGAATGTTTATATAGAGGACGAACAG	2195
Db	2808	TCTCTGCGTGGATATGTGTCTGAACTGGCTGCTGAATGTTTATATAGAGGACGAACAG	2867
Qy	2196	GAGGATCCGTGTCCGCTCTTTTAAACCTGGCATATTTCCCTGTGTAAAGACATTTGGA	2255
Db	2868	GAGGATCCGTGTCTGTCTTTTAAACCTGGCATATTTCCCTGTGTAAAGACATTTGGA	2927
Qy	2256	AGACAAATGACAGATACCTTTTCCAAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCAACG	2315
Db	2928	AGACAAATGACAGATACCTTTTCCAAAGCAAGTGGCAAGTTCAACAGGATTTTGTACCAACG	2987
Qy	2316	CAGCTGGGACCTCCTTCTGCATGATTCATCCAAATTCACAGACAGTGGGTGAAGTTGC	2375
Db	2988	CAGCTGGGACCTCCTTCTGCATGATTCATCCAAATTCACAGACAGTGGGTGAAGTTGC	3047
Qy	2376	ATCCTTTGGGGCAGTAACTTGAAC	2401
Db	3048	ATCCTTTGGGGCAGTAACTTGAAC	3073

```

RESULT 12
US-09-845-416-30
; Sequence 30, Application US/09945416
; Publication NO. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ. ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ. ID NO 30
; LENGTH: 4498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-30

```

Query Match	72.3%	Score 1737;	DB 13;	Length 4498;
Best Local Similarity	96.7%	Pred. No. 0;		
Matches 1765; Conservative	0;	Mismatches 61;	Indels 0;	Gaps 0;

Qy	576	AGATCAGCCAACTGCTGCTTTGGAGAAACAATTAAAGTATTTGGAGATCGATGGCAAA	635
Db	2005	AGAAATTAAGAAATTAAGATACAAAGACAGATGAATCTCTAAATTCAAGATGGATG	2066
Qy	636	CATCTGTAGATGAGACAGAAAGACCGCTGGCTTTTACAAGACACTCATTAATATCTGCA	695
Db	2065	CCTCAGGGTAGCTACATGAGAAAAACAAGCAATTTACATGAGACTCATTAAGTATCTCA	2124
Qy	696	ACAGTTCCTCCCTGACCTGAGAAAGATTTCTTGCTGGCTTACAGAAGCTGAACAACATGC	755
Db	2125	ACAGTTCCTCCCTGACCTGAGAAAGATTTCTTGCTGGCTTACAGAAGCTGAACAACATGC	2184

QY	756	AAAGTCCCTACAGGATCTCTACCCGTAAGAAAGGCTCCTTGAAGAATCCCAAGGAGTAATA	815
Db	2185	CAATGTCTTACAGGATCTTACCCGTAAGGAAAGGCTCTTGAAGAATCCCAAGGAGTAATA	2244
QY	816	AGAGCTGATGAAACAAATGGCAAGCTCTCAAGGTGAATTGAAGCTCACACAGATGTTTA	875
Db	2245	AGAGCTGATGAAACAAATGGCAAGCTCTCAAGGTGAATTGAAGCTCACACAGATGTTTA	2304
QY	876	TCACAACTCGATGTAATAACAGCCCAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGC	935
Db	2305	TCACAACTCGATGTAATAACAGCCCAAAAATCCTGAGATCCCTGGAAAGGTTCCGATGATGC	2364
QY	936	AGTCCTGTACAAGACGTTTGTGATAATCATGAATTCACATGAGATGAATTCGGAAAAA	995
Db	2365	AGTCCTGTACAAGACGTTTGTGATAATCATGAATTCACATGAGATGAATTCGGAAAAA	2424
QY	996	GTCTCTTACACTTTAGTTCCTTGTGAAAGCCAGTTCTGACCACTGGAAAGGCTCTGCACCT	1055
Db	2425	GTCTCTTACACTTTAGTTCCTTGTGAAAGCCAGTTCTGACCACTGGAAAGGCTCTGCACCT	2484
QY	1056	TTCTCTGACAGAACTTCTGTGTGGCTACACAGCTGAAGATGATGATTAAGACCGCAGGC	1115
Db	2485	TTCTCTGACAGAACTTCTGTGTGGCTACACAGCTGAAGATGATGATTAAGACCGCAGGC	2544
QY	1116	ACCTATTGAGAGCGACTTTCACAGCATTCAGAAACAGACGATGTACTATGGGCTTCAA	1175
Db	2545	ACCTATTGAGAGCGACTTTCACAGCATTCAGAAACAGACGATGTACTATGGGCTTCAA	2604
QY	1176	GAGGGAAATTGAAACCTAAAGAACCTGTAAATCAGAGATCTTGAAGCTGATGAAATTT	1235
Db	2605	GAGGGAAATTGAAACCTAAAGAACCTGTAAATCAGAGATCTTGAAGCTGATGAAATTT	2664
QY	1236	TCTTACAGAGACGCTTTGGAGAGCATAGAGAACTTACACAGAGGCCAAGAGCTGCC	1295
Db	2665	TCTTACAGAGAGCTTTGGAGAGCATAGAGAACTTACACAGAGGCCAAGAGCTGCC	2724
QY	1296	TCCTGAGAGAGAGCCCAAGATTGCACTCGGCTTTTACGAAAGTAGGCTGAGAGAGTCAA	1355
Db	2725	TCCTGAGAGAGAGCCCAAGATTGCACTCGGCTTTTACGAAAGTAGGCTGAGAGAGTCAA	2784
QY	1356	TACTGAGTGGGAAAAATTTGAACCTGCACCTCGCTGATCTGGCAGAGAAAAATAGATGAGAC	1415
Db	2785	TACTGAGTGGGAAAAATTTGAACCTGCACCTCGCTGATCTGGCAGAGAAAAATAGATGAGAC	2844
QY	1416	CCTTGAAGAGCTCCAGGAATTCAGAGGCCACGATGAGCTGACCTCAAGCTGCGCA	1475
Db	2845	CCTTGAAGAGCTCCAGGAATTCAGAGGCCACGATGAGCTGACCTCAAGCTGCGCA	2904
QY	1476	AGCTGAGGTATTCAGAGGATTCCTGGACCCCGTGGGGAGATCTCCTCATTTGACTCTGCCA	1535
Db	2905	AGCTGAGGTATTCAGAGGATTCCTGGACCCCGTGGGGAGATCTCCTCATTTGACTCTGCCA	2964
QY	1536	AGATCACCTCAGAGAAATCAAAGCACTTCGAGAGAAATTCGCGCTCTGAAAGAAAGT	1595
Db	2965	AGATCACCTCAGAGAAATCAAAGCACTTCGAGAGAAATTCGCGCTCTGAAAGAAAGT	3024
QY	1596	GAGCCAGTCAATGACTTTGCTCGCCAGCTTACCACTTTGGGATTTAGGCTTCAACCGTA	1655
Db	3025	GAGCCAGTCAATGACTTTGCTCGCCAGCTTACCACTTTGGGATTTAGGCTTCAACCGTA	3084
QY	1656	TAACTTCAGCACTCTGGAAGAATCGAACAACAATGAAAGTTTCTGACAGGTGGCGCTGCA	1715
Db	3085	TAACTTCAGCACTCTGGAAGAATCGAACAACAATGAAAGTTTCTGACAGGTGGCGCTGCA	3144
QY	1716	GAGCCGAGTCAAGCACTGCAATGAAGCCACAGGGAATTTGGTCAAGATTCACGACTT	1775
Db	3145	GAGCCGAGTCAAGCACTGCAATGAAGCCACAGGGAATTTGGTCAAGATTCACGACTT	3204
QY	1776	TCTTTTCAACGCTGTCCAGAGGTCCTGAGAGAGAGCCATCTCGCCAAACAAAGTGCCCTA	1835
Db	3205	TCTTTTCAACGCTGTCCAGAGGTCCTGAGAGAGAGCCATCTCGCCAAACAAAGTGCCCTA	3264
QY	1836	CTAATATCAACACAGACTCAAAACAACTTGTGTGAGCAATCCCAAAATGACAGAGCTTCA	1895

Db 3265 CTATATCAACGAGAGCTCAACCACTGCTGGGACCATCCCAAAATACAGAGCTCTA 3324
Qy 1896 CCAGCTTTAGCTGACCTGAATATATGTCAGATTTCTAGCTTTAGAGCTGCCATGAACCT 1955
Db 3325 CCAGCTTTAGCTGACCTGAATATATGTCAGATTTCTAGCTTTAGAGCTGCCATGAACCT 3384
Qy 1956 CCGAAGAGCTGACAGAGAGCCCTTGCTGATCTTGAAGCTGTCAGCTGACATGAGAGC 2015
Db 3385 CCGAAGAGCTGACAGAGAGCCCTTGCTGATCTTGAAGCTGTCAGCTGACATGAGAGC 3444
Qy 2016 CTTGACACGACACCACTTCAAGCAAAATGACACGCCATGATATCTGACAGATTATTA 2075
Db 3445 CTTGACACGACACCACTTCAAGCAAAATGACACGCCATGATATCTGACAGATTATTA 3504
Qy 2076 TTGTTGACACCAATTATGACCGCTGAGAGAGACACCAATTTGGTGAACGCTCC 2135
Db 3505 TTGTTGACCACTATTATGACCGCTGAGAGACACCAATTTGGTGAACGCTCC 3564
Qy 2136 TCTCTGCGTGAATATGTCGTGACAGCTGCTGATGTTTATGATACGGAGCAACAG 2195
Db 3565 TCTCTGCGTGAATATGTCGTGACAGCTGCTGATGTTTATGATACGGAGCAACAG 3624
Qy 2196 GAGGATCCGTGCTCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCATTTGGA 2255
Db 3625 GAGGATCCGTGCTCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCATTTGGA 3684
Qy 2256 AGACAAGTACAGATCCTTTCAAGCAAGTGGCAAGTTCAAGAGATTTTGGACCAAG 2315
Db 3685 AGACAAGTACAGATCCTTTCAAGCAAGTGGCAAGTTCAAGAGATTTTGGACCAAG 3744
Qy 2316 CAGGCTGGGCTCTCTCTGATGATTTATCCAAATTCACAGACAGTTGGTGAAGTTC 2375
Db 3745 CAGGCTGGGCTCTCTCTGATGATTTATCCAAATTCACAGACAGTTGGTGAAGTTC 3804
Qy 2376 ATCCTTTGGGGGAGAGATGATGAGC 2401
Db 3805 ATCCTTTGGGGGAGAGATGATGAGC 3830

RESULT 13

US-10-149-736-42
; Sequence 42, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-42

Query Match 72.3%; Score 1735.6; DB 13; Length 8689;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 628 TGGGCAACATCTCTGATGACAGAGACCGCTGGGTCTTTTACAGACATCATAGA 687
Db 2948 TGGAGAGCATTCATTAAGAGGTGAGTGAAGAGAGGCTGCTTGGAGAAACTCATAGA 3007

Qy 688 TTACTGCAAGATTCCCCCTGGAACCTGGAAAAAGTTCTTGCTGCTTACAGAGCTGAA 747
Db 3008 TTACTGCAAGATTCCCCCTGGAACCTGGAAAAAGTTCTTGCTGCTTACAGAGCTGAA 3067
Qy 748 ACAACCTGCAATGCTCTACAGATGCTACCCCTGATAGAAAAAGCTCTAGAAAGCTCCAG 807
Db 3068 ACAACCTGCAATGCTCTACAGATGCTACCCCTGATAGAAAAAGCTCTAGAAAGCTCCAG 3127
Qy 808 GGAATTAAGAAGCTGATGAAACCAATGGCAAGACTCCAGAGTGAATTTGAAGCTCACACA 867
Db 3128 GGAATTAAGAAGCTGATGAAACCAATGGCAAGACTCCAGAGTGAATTTGAAGCTCACACA 3187
Qy 868 GATGTTATCAACAACCTGATGAAAAACAGCCAAAAAATCTGAGATCCCTGGAGAGTTCC 927
Db 3188 GATGTTATCAACAACCTGATGAAAAACAGCCAAAAAATCTGAGATCCCTGGAGAGTTCC 3247
Qy 928 GATGATCAAGTCTCTGTTACAAAGAGCTTTGGATPAACATGAACCTTCAAGTGAAGTAACTT 987
Db 3248 GATGATCAAGTCTCTGTTACAAAGAGCTTTGGATPAACATGAACCTTCAAGTGAAGTAACTT 3307
Qy 988 CCGAAAAAGTCTGCAACAATTAGTCCATTGGAAAGCCAGTCTGACCACTGGAAAGCGT 1047
Db 3308 CCGAAAAAGTCTGCAACAATTAGTCCATTGGAAAGCCAGTCTGACCACTGGAAAGCGT 3367
Qy 1048 CTGCACTTTCTCTGACAGAACTTCTGTGTGCTACAGCTGAAAGATGATGATTAAGC 1107
Db 3368 CTGCACTTTCTCTGACAGAACTTCTGTGTGCTACAGCTGAAAGATGATGATTAAGC 3427
Qy 1108 CCGCAGGACCTTATTTGAGAGGCACTTTCCAGCAGTTCCAGAGCAAGACGATGATTAAG 1167
Db 3428 CCGCAGGACCTTATTTGAGAGGCACTTTCCAGCAGTTCCAGAGCAAGACGATGATTAAG 3487
Qy 1168 GCCTTCAAGAGGAAATTGAAAACTAAGAACTGATATCATGATGATCTTGAAGCTGTA 1227
Db 3488 GCCTTCAAGAGGAAATTGAAAACTAAGAACTGATATCATGATGATCTTGAAGCTGTA 3547
Qy 1228 CGAATATTTTGAACAAGACGCTTTGAAAGGACTAGAGAACTCTTCAAGAGCCCAAG 1287
Db 3548 CGAATATTTTGAACAAGACGCTTTGAAAGGACTAGAGAACTCTTCAAGAGCCCAAG 3607
Qy 1288 GAGCTGCTCTGAGAGAGAGGCCAGAAATGCTACGCTGCTTCAAGAAAGCGCTGAG 1347
Db 3608 GAGCTGCTCTCTGAGAGAGAGGCCAGAAATGCTACGCTGCTTCAAGAAAGCGCTGAG 3667
Qy 1348 GAGGTCAATACCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACCTGGAGAGAAAAATA 1407
Db 3668 GAGGTCAATACCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACCTGGAGAGAAAAATA 3727
Qy 1408 GATGAGACCTTTAAAGACTCCAGAGAACTTCAAGAGGCCAGGATGAGCTGAGACCTCAAG 1467
Db 3728 GATGAGACCTTTAAAGACTCCAGAGAACTTCAAGAGGCCAGGATGAGCTGAGACCTCAAG 3787
Qy 1468 CTGCGCAAGCTGAGGTGATCAAGAGGATCTTGAGAGCCGCTGGAGCATCTCTCATTTGAGC 1527
Db 3788 CTGCGCAAGCTGAGGTGATCAAGAGGATCTTGAGAGCCGCTGGAGCATCTCTCATTTGAGC 3847
Qy 1528 TCTTCTCAAGATACCTCGAAGAAAGTCAAGGCACTTGAAGAGAAATTCGCGCTCTGAAA 1587
Db 3848 TCTTCTCAAGATACCTCGAAGAAAGTCAAGGCACTTGAAGAGAAATTCGCGCTCTGAAA 3907
Qy 1588 GAGAAAGTGAAGCAAGCTCAATGACCTTGCTGCGAGCTTACACACTTTGGGCAATTGAGCTC 1647
Db 3908 GAGAAAGTGAAGCAAGCTCAATGACCTTGCTGCGAGCTTACACACTTTGGGCAATTGAGCTC 3967
Qy 1648 TCACGTTATACCTGACACTCTGGAAGACCTGGAACCAACAGATGGAAGCTTCTGACAGTG 1707
Db 3968 TCACGTTATACCTGACACTCTGGAAGACCTGGAACCAACAGATGGAAGCTTCTGACAGTG 4027
Qy 1708 GCGGTGAGAGAACGATGACAGGAGCTGATGAAGCCCAAGAGGACTTGGTCCAGAGATCT 1767
Db 4028 GCGGTGAGAGAACGATGACAGGAGCTGATGAAGCCCAAGAGGACTTGGTCCAGAGATCT 4087
Qy 1768 CAGCACTTTCTTTCCACGCTGTGTCAGAGGCTCCTGGAGAGAGCCATCTGCCAAACAAA 1827

Db 4088 CAGCATTCTTCTTCCACGCTGTCCAGGGTCCCTGGAGAGAGCCATCTCCCAAAACAA 4147
Qy 1828 GTGCCCTATATATCAACAGAGCTCAAACTTGTCTGGAGACATCCCAATGACA 1887
Db 4148 GTGCCCTATATATCAACAGAGCTCAAACTTGTCTGGAGACATCCCAATGACA 4287
Qy 1888 GAGCTCTACAGCTCTTTAGCTGACCTGAATAATGTCAATTTCTCACTTATAGACTGCC 1947
Db 4208 GAGCTCTACAGCTCTTTAGCTGACCTGAATAATGTCAATTTCTCACTTATAGACTGCC 4267
Qy 1948 AAGAACTCCGAGAGCTGAGAGAGGCCCTTTGCTGTGATCTCTTGAAGCTCTGAGTGA 2007
Db 4268 ATGAACTCCGAGAGCTGAGAGAGGCCCTTTGCTGTGATCTCTTGAAGCTCTGAGTGA 4327
Qy 2008 TGTGATGCTTGGACCAACCACTCAAGCAAAATGACAGCCCATGATATCTGAG 2067
Db 4328 TGTGATGCTTGGACCAACCACTCAAGCAAAATGACAGCCCATGATATCTGAG 4387
Qy 2068 ATTATTAATGTTTGAACAATTTATGACCGCTTGGAGCAAGAGCAACAATTTGCTC 2127
Db 4388 ATTATTAATGTTTGAACAATTTATGACCGCTTGGAGCAAGAGCAACAATTTGCTC 4447
Qy 2128 AACGTCCTCTCTGCTGATATATGTCTGAACTGCTGATGTTTATGATACGGA 2187
Db 4448 AACGTCCTCTCTGCTGATATATGTCTGAACTGCTGATGTTTATGATACGGA 4507
Qy 2188 CGAACAGGAGAGATCCGTCTCTCTTTTAAATGCGATCATTTCCCTGTGTAAGCA 2247
Db 4508 CGAACAGGAGAGATCCGTCTCTCTTTTAAATGCGATCATTTCCCTGTGTAAGCA 4567
Qy 2248 CATTGGAAGACAATACAGATACCTTTTCAAGCAAGGCAAGTTCACAGATTTGT 2307
Db 4568 CATTGGAAGACAATACAGATACCTTTTCAAGCAAGGCAAGTTCACAGATTTGT 4627
Qy 2308 GACCAAGCGAGGCTGGGCTCTCTTCTGATGATTTTATCAAAATTCAGACAGTTGGT 2367
Db 4628 GACCAAGCGAGGCTGGGCTCTCTTCTGATGATTTTATCAAAATTCAGACAGTTGGT 4687
Qy 2368 GAAATTGATCTCTTTGGGGGAGTAACTTGGC 2401
Db 4688 GAAATTGATCTCTTTGGGGGAGTAACTTGGC 4721

RESULT 14
US-09-845-416-1
; Sequence 1, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DELI42
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-1

Query Match 72.3%; Score 1735.6; DB 13; Length 11058;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 628 TGGGCAAAATCTCTAGATGAGACAGAGCCGCTGGTCTTTTCAAGACATCATAGA 687
Db 8008 TGGGCAAAATCTCTAGATGAGACAGAGCCGCTGGTCTTTTCAAGACATCATAGA 8067

Qy 688 TTACTGCAACAGTTCCTCCCTGAGCCTGGAAAAGTTTCTTGTGCTTACAGAGCTGAA 747
Db 8068 TTACTGCAACAGTTCCTCCCTGAGCCTGGAAAAGTTTCTTGTGCTTACAGAGCTGAA 8127
Qy 748 ACACTGGCAATGCTCTACAGATGTACCGTAAAGAAAGGCTCCTGAAGACCTCAAG 807
Db 8128 ACACTGGCAATGCTCTACAGATGTACCGTAAAGAAAGGCTCCTGAAGACCTCAAG 8187
Qy 808 GGAATAAAAGCTGATGAAACAAATGCAAGACCTTCAAGGTGAATTTGAAGCTCAACA 867
Db 8188 GGAATAAAAGCTGATGAAACAAATGCAAGACCTTCAAGGTGAATTTGAAGCTCAACA 8247
Qy 868 GATGTTATCAACCTGATGAAACACCCAAAATTCCTGATGATCCCTGGAAGTTC 927
Db 8248 GATGTTATCAACCTGATGAAACACCCAAAATTCCTGATGATCCCTGGAAGTTC 8307
Qy 928 GATGATGAGTCTGTTACAAAGAGTGTGATTAATGATGATGATGATGATGATGAT 987
Db 8308 GATGATGAGTCTGTTACAAAGAGTGTGATTAATGATGATGATGATGATGATGAT 8367
Qy 988 CGGAAAAAGTCTCTCAACATTAAGTCCCTTTGGAAGCCAGTTCGACCAAGTGAAGGT 1047
Db 8368 CGGAAAAAGTCTCTCAACATTAAGTCCCTTTGGAAGCCAGTTCGACCAAGTGAAGGT 8427
Qy 1048 CTGCACTTCTCTGAGAGAACTCTGATGAGTCAAGCTGAAGATGATGATGATGATGAT 1107
Db 8428 CTGCACTTCTCTGAGAGAACTCTGATGAGTCAAGCTGAAGATGATGATGATGATGAT 8487
Qy 1108 CGGAGGCACTTATGAGAGCACTTTTCAAGAGTTCAGAGCAAGAGATGATGATGATGAT 1167
Db 8488 CGGAGGCACTTATGAGAGCACTTTTCAAGAGTTCAGAGCAAGAGATGATGATGATGAT 8547
Qy 1168 GCCTCAAGAGGAAATGAAAATTAAGAACTGATATCATGATGATGATGATGATGATGAT 1227
Db 8548 GCCTCAAGAGGAAATGAAAATTAAGAACTGATATCATGATGATGATGATGATGATGAT 8607
Qy 1228 CGAATATTTCTGACAGAGAGGCTTTGAAAGACTTGAAGAACTTACAGAGAGCCACA 1287
Db 8608 CGAATATTTCTGACAGAGAGGCTTTGAAAGACTTGAAGAACTTACAGAGAGCCACA 8667
Qy 1288 GAGTGCCTCTGAGAGAGAGCCAGAAATGTCATCTGCTCTGCAAGAGAGGCTGAG 1347
Db 8668 GAGTGCCTCTGAGAGAGAGCCAGAAATGTCATCTGCTCTGCAAGAGAGGCTGAG 8727
Qy 1348 GAGTCAATCTGAGTGGGAAAATTTGAACCTGCACTCGCTGACCTGCAAGAAAAATA 1407
Db 8728 GAGTCAATCTGAGTGGGAAAATTTGAACCTGCACTCGCTGACCTGCAAGAAAAATA 8787
Qy 1408 GATGAAACCTTGAAGAACTCCAGAACTTCAAGAGCCACGAGATGAGTGAACCTCAAG 1467
Db 8788 GATGAAACCTTGAAGAACTCCAGAACTTCAAGAGCCACGAGATGAGTGAACCTCAAG 8847
Qy 1468 CTGCGCAAGCTGAGTGAATCAAGGATCTGCGAGCCGATGGAGATCTCTCATTTAGC 1527
Db 8848 CTGCGCAAGCTGAGTGAATCAAGGATCTGCGAGCCGATGGAGATCTCTCATTTAGC 8907
Qy 1528 TCTCTCAAGATCACTGAGAAAGTCAAGGCACTTCAAGAGAAATTTGCGCTTGA 1587
Db 8908 TCTCTCAAGATCACTGAGAAAGTCAAGGCACTTCAAGAGAAATTTGCGCTTGA 8967
Qy 1588 GAGAACGTAGAGCACTCAATGACCTGCTGCGCAAGCTTACAGCTTTGGGCAATGAGCTC 1647
Db 8968 GAGAACGTAGAGCACTCAATGACCTGCTGCGCAAGCTTACAGCTTTGGGCAATGAGCTC 9027
Qy 1648 TCACGATTAACCTGACAGCTCTGGAAGCTTGAACACCAAGATGAAGTTCTTGAAGG 9087
Db 9028 TCACGATTAACCTGACAGCTCTGGAAGCTTGAACACCAAGATGAAGTTCTTGAAGG 9147
Qy 1708 GCGGTGAGAGCCAGTCAAGGAGCTGATGAAGCCACAGGAGCTTTGTCAGATCT 1767
Db 9088 GCGGTGAGAGCCAGTCAAGGAGCTGATGAAGCCACAGGAGCTTTGTCAGATCT 9147
Qy 1768 CAGCACTTTTCTTCAACGCTGTCTGCAAGGCTCTGGAGAGAGCACTCTGCAAAACAA 1827

Db 9148 CAGCACTTCTTCCACGCTCTGTCAGGGTCCCTGGAGAGAGCCATCTCGCAACAA 9207
Qy 1828 GTGCCCTACTATATCAACACAGACTCAACAACTGCTGGACCATCCCAAAATGCA 1887
Db 9208 GTGCCCTACTATATCAACACAGACTCAACAACTGCTGGACCATCCCAAAATGCA 9267
Qy 1888 GAGCTCTACAGCTCTTACCTGACCTGAAATATGTCAGATTCAGCTTATGAGCTGCC 1947
Db 9268 GAGCTCTACAGCTCTTACCTGACCTGAAATATGTCAGATTCAGCTTATGAGCTGCC 9327
Qy 1948 ATGAACTCCGAAAGCTGAGAAAGGCCCTTCTGAGATCTCTGAGCTGTCAGCTCA 2007
Db 9328 ATGAACTCCGAAAGCTGAGAAAGGCCCTTCTGAGATCTCTGAGCTGTCAGCTCA 9387
Qy 2008 TGTGATGCTTGGACACAGACAACTCAAGCAAAATGACACCACTGATTCCTGAG 2067
Db 9388 TGTGATGCTTGGACACAGACAACTCAAGCAAAATGACACCACTGATTCCTGAG 9447
Qy 2068 ATTATTAATGTTGACACATAATTATGACCCCTGAGACAGACACAACTTGGTC 2127
Db 9448 ATTATTAATGTTGACACATAATTATGACCCCTGAGACAGACACAACTTGGTC 9507
Qy 2128 AACGTCCCTCTGCTGCTGATATGTCCTGAACTGCTGATGTTATGATTCGGA 2187
Db 9508 AACGTCCCTCTGCTGCTGATATGTCCTGAACTGCTGATGTTATGATTCGGA 9567
Qy 2188 CGAACAGAGAGAGTCCGTCTCTGTTTAAAGCTGGATCATTTCCCTGTGTAACA 2247
Db 9568 CGAACAGAGAGATCCGTCTCTGTTTAAAGCTGGATCATTTCCCTGTGTAACA 9627
Qy 2248 CATTGGAAGACAGTACAGATACCTTTCAAGCAAGTGGACAGTTCAACAGATTTGT 2307
Db 9628 CATTGGAAGACAGTACAGATACCTTTCAAGCAAGTGGACAGATTTGT 9687
Qy 2308 GACCAAGCGAGCTGGGCTCTCTGATGATTTATTCAAATTCGAAGAGTGGGT 2367
Db 9688 GACCAAGCGAGCTGGGCTCTCTGATGATTTATTCAAATTCGAAGAGTGGGT 9747
Qy 2368 GAGTTGATCTCTTGGGGGAGTAACTTGAGC 2401
Db 9748 GAGTTGATCTCTTGGGGGAGTAACTTGAGC 9781

RESULT 15
US-10-149-736-44
; Sequence 44, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 11443
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-44

Query Match 72.3%; Score 1735.6; DB 13; Length 11443;
Beet Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 628 TGGCAACATCTGTAGATGAGACAGAACCCGCTGGTCTTTTACAGACACTCATAGA 687
Db 5702 TGGAGAGCATCATATAAAAGGTGAGTGAACGAGAGCGCTTGGAAAGCACTCATAGA 5761
Qy 688 TTACTGCAAGATTTCCCTGAGCTGAGAAAGTTTCTGCTGCTGCTGCTGAGAGCTGAA 747
Db 5762 TTACTGCAAGATTTCCCTGAGCTGAGAAAGTTTCTGCTGCTGCTGAGAGCTGAA 5821
Qy 748 ACAACTGCCAATGCTCTACAGATGCTACCCCTGAGAAAGGCTCTGAGAGCTCAAG 807
Db 5822 ACAACTGCCAATGCTCTACAGATGCTACCCCTGAGAAAGGCTCTGAGAGCTCAAG 5881
Qy 808 GAGTAAAGAGCTGATGAAACATGAGCAAGACCTCAAGGTGAATTTGAGGTCAACA 867
Db 5882 GAGTAAAGAGCTGATGAAACATGAGCAAGACCTCAAGGTGAATTTGAGGTCAACA 5941
Qy 868 GATGTTTATCAAACTGTGATGAAACAGCCAAATAATCCTGAGATCCCTGAGAGTTC 927
Db 5942 GATGTTTATCAAACTGTGATGAAACAGCCAAATAATCCTGAGATCCCTGAGAGTTC 6001
Qy 928 GATGATGCACTCTGTTTACAAAGAGCTTTGATGATTAATGATCACTTCAAGGTGAACTT 987
Db 6002 GATGATGCACTCTGTTTACAAAGAGCTTTGATGATTAATGATCACTTCAAGGTGAACTT 6061
Qy 988 CGAATAAAGTCTCTCAACATTAAGTCCATTTGAGACAGATTCAGCAGTGAAGCGT 1047
Db 6062 CGAATAAAGTCTCTCAACATTAAGTCCATTTGAGACAGATTCAGCAGTGAAGCGT 6121
Qy 1048 CTGACCTTTCTCTGACAGAACTTCTGTGTGCTACAGCTGAAAGATGATGATTAAGC 1107
Db 6122 CTGACCTTTCTCTGACAGAACTTCTGTGTGCTACAGCTGAAAGATGATGATTAAGC 6181
Qy 1108 CGGACAGACCTATTTGAGAGCGCACTTCCAGAGTTCAGAAAGACAGATGATCATAG 1167
Db 6182 CGGACAGACCTATTTGAGAGCGCACTTCCAGAGTTCAGAAAGACAGATGATCATAG 6241
Qy 1168 GCCTTCAAGAGGGAATGAAACTTAAGAACCTGATATCATGATGATCTTTGAGACTGTA 1227
Db 6242 GCCTTCAAGAGGGAATGAAACTTAAGAACCTGATATCATGATGATCTTTGAGACTGTA 6301
Qy 1228 CGAATATTTCTGACAGAGAGCTTTGAGAGGCTGAGAAACTCTACAGAGAGCCCA 1287
Db 6302 CGAATATTTCTGACAGAGAGCTTTGAGAGGCTGAGAAACTCTACAGAGAGCCCA 6361
Qy 1288 GAGCTGCTCTCTGAGAGAGAGCCAGATGTCATCTGCTTCTACGAAGAGGCTGAG 1347
Db 6362 GAGCTGCTCTCTGAGAGAGAGCCAGATGTCATCTGCTTCTACGAAGAGGCTGAG 6421
Qy 1348 GAGGTCAATGATGAGGGAATTTGAACCTGCACTCGCTGATCTGGCAGGAAATA 1407
Db 6422 GAGGTCAATGATGAGGGAATTTGAACCTGCACTCGCTGATCTGGCAGGAAATA 6481
Qy 1408 GATGAGACCTTGAAGACCTCAAGAACTTCAAGAGGCGCAGATGATGAGACCTCAAG 1467
Db 6482 GATGAGACCTTGAAGACCTCAAGAACTTCAAGAGGCGCAGATGATGAGACCTCAAG 6541
Qy 1468 CTGCGCAAGCTGAGGTGATCAAGGATCTGAGAGCCGCTGGCGATCTCTCATTTGAC 1527
Db 6542 CTGCGCAAGCTGAGGTGATCAAGGATCTGAGAGCCGCTGGCGATCTCTCATTTGAC 6601
Qy 1528 TCTCTCCAGATCACTCGAGAAAGTCAAGGACTTCAGAGGAATTTGGGCTCTGAA 1587
Db 6602 TCTCTCCAGATCACTCGAGAAAGTCAAGGACTTCAGAGGAATTTGGGCTCTGAA 6661
Qy 1588 GAGACGTGAGCAAGTCAATGACCTTGTCTGCGCAGCTTACCACTTTGGGCAATGAGCTTC 1647
Db 6662 GAGACGTGAGCAAGTCAATGACCTTGTCTGCGCAGCTTACCACTTTGGGCAATGAGCTTC 6721
Qy 1648 TACCGTAACTCAAGCACTTGTGAAAGCTGAACTCAAGATGATGAGGCTTGTGAGG 1707
Db 6722 TACCGTAACTCAAGCACTTGTGAAAGCTGAACTCAAGATGATGAGGCTTGTGAGG 6781
Qy 1708 GCGGTGAGAGAGAGTCAAGGAGCTGATGAGGCCACAGGACTTTGTGTCAGACATCT 1767

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 1, 2004, 11:11:40 ; Search time 148.424 Seconds
(without alignments)
7140.092 Million cell updates/sec

Title: US-09-845-416-9_COPY_1000_3400
Perfect score: 2401

Sequence: 1 ggcagctcatcgatcgagagag.....tgggggcagctcatcgagc 2401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 56978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents NA: *
1: /cgn2_6/pdata/2/ina/5A.COMB.seq: *
2: /cgn2_6/pdata/2/ina/5B.COMB.seq: *
3: /cgn2_6/pdata/2/ina/6A.COMB.seq: *
4: /cgn2_6/pdata/2/ina/6B.COMB.seq: *
5: /cgn2_6/pdata/2/ina/6C.COMB.seq: *
6: /cgn2_6/pdata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1735.6	72.3	5952	US-09-687-875A-1	Sequence 1, Appli
2	1724.6	71.8	13977	US-09-484-970B-60	Sequence 60, Appli
3	1506.8	62.8	19307	US-08-836-022A-10	Sequence 10, Appli
4	1506.8	62.8	19307	US-09-427-048A-10	Sequence 10, Appli
5	690.6	28.8	6045	US-09-091-501B-7	Sequence 7, Appli
6	690.6	28.8	10320	US-09-091-501B-9	Sequence 9, Appli
7	79.4	3.3	200	US-09-091-501B-4	Sequence 4, Appli
8	78.6	3.3	200	US-09-091-501B-6	Sequence 6, Appli
9	78.6	3.3	200	US-09-091-501B-14	Sequence 14, Appli
10	76.6	3.2	7218	US-09-687-875A-13	Sequence 13, Appli
11	63.6	2.6	238	US-09-668-312A-10	Sequence 10, Appli
12	44.2	1.8	2574	US-09-198-452A-1	Sequence 1, Appli
13	44	1.8	1230025	US-09-107-532A-1186	Sequence 1186, Appli
14	43.4	1.8	1179	US-09-620-312D-69	Sequence 69, Appli
15	42.8	1.8	1590	US-08-368-590-1	Sequence 1, Appli
16	42.8	1.7	7812	US-08-257-073-4	Sequence 4, Appli
17	40.4	1.7	2223	US-08-961-527-82	Sequence 82, Appli
18	39.2	1.6	16995	US-09-620-312D-847	Sequence 847, Appli
19	38.6	1.6	1751	US-08-425-069-3	Sequence 3, Appli
20	38.6	1.6	1995	US-08-317-844B-3	Sequence 24, Appli
21	38.6	1.6	1995	US-09-220-132-24	Sequence 3, Appli
22	38.4	1.6	7672	US-09-668-312A-3	Sequence 3, Appli
23	38.2	1.6	428	US-09-668-312A-3	Sequence 3, Appli
24	38.2	1.6	1131	US-09-668-312A-3	Sequence 3, Appli
25	38.2	1.6	1784	US-09-668-312A-3	Sequence 3, Appli
26	38.2	1.6	4439	US-09-668-312A-3	Sequence 3, Appli
27	38	1.6	1394	US-09-247-155-76	Sequence 76, Appli

28	36.8	1.5	1886	6	5210183-1	Patent No. 5210183
29	36.6	1.5	1845	4	US-08-887-534A-22	Sequence 22, Appli
30	36.6	1.5	1845	4	US-09-527-431-22	Sequence 22, Appli
31	36.2	1.5	289	3	US-09-007-005-17	Sequence 17, Appli
32	36.2	1.5	289	3	US-09-244-796-17	Sequence 17, Appli
33	36.2	1.5	2447	2	US-09-014-969-14	Sequence 14, Appli
34	36.2	1.5	16875	4	US-09-426-290-1	Sequence 1, Appli
35	36	1.5	2873	4	US-08-630-915A-193	Sequence 193, Appli
36	35.8	1.5	790	3	US-09-461-474-13	Sequence 13, Appli
37	35.6	1.5	4868	1	US-08-139-937-12	Sequence 12, Appli
38	35.6	1.5	4868	5	PCT-US93-11310-12	Sequence 12, Appli
39	35.6	1.5	8257	4	US-09-595-684B-30	Sequence 30, Appli
40	35.6	1.5	8789	1	US-08-328-254-5	Sequence 5, Appli
41	35.6	1.5	10136	1	US-08-353-700-2	Sequence 2, Appli
42	35.6	1.5	10136	5	PCT-US95-16216-2	Sequence 2, Appli
43	35.2	1.5	3466	1	US-08-468-036-38	Sequence 38, Appli
44	35.2	1.5	3466	2	US-08-376-843-38	Sequence 38, Appli
45	35	1.5	2169	4	US-09-434-408-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-687-875A-1
Sequence 1, Application US/09687875A
Patent No. 6544786
GENERAL INFORMATION:
APPLICANT: Xiao, Xiao
APPLICANT: Liu, Paul
TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICEI
FILE REFERENCE: 00792
CURRENT APPLICATION NUMBER: US/09/687,875A
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/158,868
PRIOR FILING DATE: 1999-10-15
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 5952
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (2897)..(3198)
OTHER INFORMATION: S4 junction site
NAME/KEY: misc feature
LOCATION: (3198)..(3199)
OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match

Best Local Similarity 72.3%; Score 1735.6; DB 4; Length 5952;

Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY	628	TGGGCAACATCTGTATGATGACAGAACCCGCTGCTTTTACAAACACTATAA	687
DB	2902	TGAGAGACATTCATATAAGGATGATGACAGAGGCTCTTGAAGAACTCATACA	2961
QY	688	TTACTGCAACAGTTCCCTCGACCTGCAAGAAAGTTCTTCTGCTGCTTACAGAGTGA	747
DB	2962	TTACTGCAACAGTTCCCTCGACCTGCAAGAAAGTTCTTCTGCTGCTTACAGAGTGA	3021
QY	748	ACAACATGCAATGCTTACAGATGATGACAGTACCCGTAAGAAAGCTCTACAGACTCAAG	807
DB	3022	ACAACATGCAATGCTTACAGATGATGACAGTACCCGTAAGAAAGCTCTACAGACTCAAG	3081
QY	808	GAGATTAAGAGGTGATGAAGCAATGCAAGACCTCCAGAGGAATTGAAGTACACAA	867
DB	3082	GAGATTAAGAGGTGATGAAGCAATGCAAGACCTCCAGAGGAATTGAAGTACACAA	3141
QY	868	GATGTTATCAACACTGATGATAAACAACCAAAATCTTGATGCTCTGAAAGTTCC	927

Db 3142 GATGTTATCAACAACCTGATGAAAGCAAGCCAAAATCTGAGATCCCTGAGAGTTCC 3201
 Qy 928 GATATGCAAGTCTCTTACAAAGAGCTTTGATTAACATGAACCTCAAGTGGATGAACCTT 987
 Db 3202 GATATGCAAGTCTCTTACAAAGAGCTTTGATTAACATGAACCTCAAGTGGATGAACCTT 3261
 Qy 988 CGAAGAAAGTCTCAACATAGTCCCATTTGGAGCCAGTTCTGACCAAGTGAAGCGT 1047
 Db 3262 CGAAGAAAGTCTCTCAACATAGTCCCATTTGGAGCCAGTTCTGACCAAGTGAAGCGT 3321
 Qy 1048 CTGCACTTTCTCTGCAAGAACTTCTGTGTGCTTAACGCTGAAGATGATGAATTAAC 1107
 Db 3322 CTGCACTTTCTCTGCAAGAACTTCTGTGTGCTTAACGCTGAAGATGATGAATTAAGC 3381
 Qy 1108 CGGCAAGCACTTATTTGAGAGCGACTTTCAGCACTTGAAGAGCAAGATGATGAAG 1167
 Db 3382 CGGCAAGCACTTATTTGAGAGCGACTTTCAGCACTTGAAGAGCAAGATGATGAAG 3441
 Qy 1168 GCTTTCAAGAGGAAATGAAATCTAAAGACCTGTATCATGAGTACTCTTGAAGCTGA 1227
 Db 3442 GCTTTCAAGAGGAAATGAAATCTAAAGACCTGTATCATGAGTACTCTTGAAGCTGA 3501
 Qy 1228 CGAATATTTCTGACAGAGCAAGCTTTGAGAGCACTAAGAAACCTTACCGAGCCCGAG 1287
 Db 3502 CGAATATTTCTGACAGAGCAAGCTTTGAGAGCACTAAGAAACCTTACCGAGCCCGAG 3561
 Qy 1288 GAGTGCTCTCTGAGAGAGAGCCAGATGCTCACTGCTTCTACGAAAGAGAGCTGAG 1347
 Db 3562 GAGTGCTCTCTGAGAGAGAGCCAGATGCTCACTGCTTCTACGAAAGAGAGCTGAG 3621
 Qy 1348 GAGTCAATACTGAGTGGGAAATTTGAACCTGCACTCCGCTGACTGCGAGAGAAATTA 1407
 Db 3622 GAGTCAATACTGAGTGGGAAATTTGAACCTGCACTCCGCTGACTGCGAGAGAAATTA 3681
 Qy 1408 GATGAGACCTTTGAAAGAGCTCCAGAACTTCAAGAGCCAGAGTGAAGTGAAGCTTCAAG 1467
 Db 3682 GATGAGACCTTTGAAAGAGCTCCAGAACTTCAAGAGCCAGAGTGAAGTGAAGCTTCAAG 3741
 Qy 1468 CTGGGCAAGCTGAGAGTGAATCAAGGATCTGAGAGCCGCTGAGCGATCTTCTCATTGAC 1527
 Db 3742 CTGGGCAAGCTGAGAGTGAATCAAGGATCTGAGAGCCGCTGAGCGATCTTCTCATTGAC 3801
 Qy 1528 TCTCTCAAGATCACTGAGAAAGTCAAGGCACTTGAAGAGAAATTTGGCTCTGAAA 1587
 Db 3802 TCTCTCAAGATCACTGAGAAAGTCAAGGCACTTGAAGAGAAATTTGGCTCTGAAA 3861
 Qy 1588 GAGAACTGAGCCAGCTCAATGATGACTTCTGAGAGTCAAGCTTCAAGCTTCAAGCTC 1647
 Db 3862 GAGAACTGAGCCAGCTCAATGATGACTTCTGAGAGTCAAGCTTCAAGCTTCAAGCTC 3921
 Qy 1648 TCACCGTAACTCAGACACTTGGAGAGCCTGAACCCAGATGGAAGCTTTCGAGAGTG 1707
 Db 3922 TCACCGTAACTCAGACACTTGGAGAGCCTGAACCCAGATGGAAGCTTTCGAGAGTG 3981
 Qy 1708 GCCGTGAGAGCCAGCTGAGAGCTGATGAGCCAGAGGCACTTTGATTCAGACATCT 1767
 Db 3982 GCCGTGAGAGCCAGCTGAGAGCTGATGAGCCAGAGGCACTTTGATTCAGACATCT 4041
 Qy 1768 CAGCACTTTCTTCCAGCTGTCTGAGAGGCTCCGAGAGAGAGCACTTCTGCAAAACAA 1827
 Db 4042 CAGCACTTTCTTCCAGCTGTCTGAGAGGCTCCGAGAGAGAGCACTTCTGCAAAACAA 4101
 Qy 1828 GTGCCCTACTATCAACAGAGACTCAACCACTTGTGAGAGCACTTCCCAAAATGACA 1887
 Db 4102 GTGCCCTACTATCAACAGAGACTCAACCACTTGTGAGAGCACTTCCCAAAATGACA 4161
 Qy 1888 GAGCTTACAGCTTTTACCTGAGAGCTGATATGATGATTTCACTTTATAGAGCTGCC 1947
 Db 4162 GAGCTTACAGCTTTTACCTGAGAGCTGATATGATGATTTCACTTTATAGAGCTGCC 4221
 Qy 1948 ATGAACTCGAGAGCTGAGAGAGCTTGTGATGATCTTGTGAGAGCTGAGCTGCA 2007
 Db 4222 ATGAACTCGAGAGCTGAGAGAGCTTGTGATGATCTTGTGAGAGCTGAGCTGCA 4281

Qy 2008 TGATGCTTGGACCAAGACACACCTCAAGCAAAATGACGAGCCCATGATATCTGAG 2067
 Db 4282 TGATGCTTGGACCAAGACACACCTCAAGCAAAATGACGAGCCCATGATATCTGAG 4341
 Qy 2068 ATTATTAATTGTTTACCACTATTATGACCGCTGAGAGCAAGCAACAAATTTGCTC 2127
 Db 4342 ATTATTAATTGTTTACCACTATTATGACCGCTGAGAGCAAGCAACAAATTTGCTC 4401
 Qy 2128 AACGTCCTCTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2187
 Db 4402 AACGTCCTCTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4461
 Qy 2188 CGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2247
 Db 4462 CGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4521
 Qy 2248 CATTGGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2307
 Db 4522 CATTGGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4581
 Qy 2308 GAGCAGCGAGAGCTGAGCTCTCTCTGATGATGATGATGATGATGATGATGATGATG 2367
 Db 4582 GAGCAGCGAGAGCTGAGCTCTCTCTGATGATGATGATGATGATGATGATGATGATG 4641
 Qy 2368 GAGTTCATCTCTTGGGAGAGTAACTTGAAC 2401
 Db 4642 GAGTTCATCTCTTGGGAGAGTAACTTGAAC 4675

RESULT 2
 US-09-484-970B-60
 / Sequence 60, Application US/09484970B
 / Patent No. 6426186
 / GENERAL INFORMATION:
 / APPLICANT: Jones, Karen A.
 / APPLICANT: Volkmuth, Wayne
 / APPLICANT: Walker, Michael G.
 / TITLE OF INVENTION: BONE REMODELING GENES
 / FILE REFERENCE: PB-0014 US
 / CURRENT APPLICATION NUMBER: US/09/484,970B
 / NUMBER OF SEQ ID NOS: 2000-01-18
 / SOFTWARE: PERL Program
 / SEQ ID NO. 60
 / LENGTH: 13977
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: misc feature
 / OTHER INFORMATION: Incyte ID No. 6426186 229357.11CBI
 / NAME/KEY: unsure
 / LOCATION: 11721-11761, 12294, 13969
 / OTHER INFORMATION: a, t, c, g, or other
 US-09-484-970B-60

Query Match 71.8%; Score 1724.6; DB 4; Length 13977;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1150; Conservative 0; Mismatches 24; Indels 1; Gaps 1;
 Qy 628 TGGGCAAAACATTTGATGATGACAGAGAGCCGCTGAGTCTTTTCAAGACATCATGGA 687
 Db 8216 TGGAAACACATTTCAATTAAGAGGATGATGAGAGAGAGGCTGTTTGAAGAAATCTCATGGA 8275
 Qy 688 TTACTGCAACAGTTCCCTCGAGACCTGGAAGAAAGTTTCTTGCCTGAGCTTACAGAGCTGAA 747
 Db 8276 TTACTGCAACAGTTCCCTCGAGACCTGGAAGAAAGTTTCTTGCCTGAGCTTACAGAGCTGAA 8335
 Qy 748 ACAACTGCCAATGCTTCAAGAGTCTACCCGTAAGAAAGGCTCTTGAAGACTCCAG 807
 Db 8336 ACAACTGCCAATGCTTCAAGAGTCTACCCGTAAGAAAGGCTCTTGAAGAGACTCCAG 8395
 Qy 808 GGAATTAAGAGCTGATGAACCAATGCAAGACCTCCAAAGGTGAATGAAGTCAACACA 867

DB 8396 GAGGTAAGAGCTGATGAAACATGCGAAGCTCCGAAAGTGAATGAAAGCTCAGCA 8455
QY 868 GATGTTTATCACACCTGGATGAAACAGCCAAAAATCTGAGATCCCTGGAGTTCC 927
DB 8456 GATGTTTATCACACCTGGATGAAACAGCCAAAAATCTGAGATCCCTGGAGTTCC 8515
QY 928 GATGATGAGTCTCTGTTACAAAGAGTTTGAATACATGAATTTCAAGTGAATTT 987
DB 8516 GATGATGAGTCTCTGTTACAAAGAGTTTGAATACATGAATTTCAAGTGAATTT 8575
QY 988 CGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCAGTTGACAGAGTGAAGGT 1047
DB 8576 CGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCAGTTGACAGAGTGAAGGT 8635
QY 1048 CTGACCTTTCTCTGAGAGAACTTCTGTGTGGCTACAGCTGAAAGATGAATTAAGC 1107
DB 8636 CTGACCTTTCTCTGAGAGAACTTCTGTGTGGCTACAGCTGAAAGATGAATTAAGC 8695
QY 1108 CGGCAAGGACCTATTGAGAGGAGCTTCCAGAGTTGAGAGCAAGATGATCATAG 1167
DB 8696 CGGCAAGGACCTATTGAGAGGAGCTTCCAGAGTTGAGAGCAAGATGATCATAG 8755
QY 1168 GCCTTCAGAGGGAATTGAAAACTAAAGAACCTGTATCATGATGACTCTTGAGACTGTA 1227
DB 8756 GCCTTCAGAGGGAATTGAAAACTAAAGAACCTGTATCATGATGACTCTTGAGACTGTA 8815
QY 1228 CGAATATTTTCTGACAGAGAGCTTTGGAAGACTGAGAAACTTTACAGAGAGCCGCA 1287
DB 8816 CGAATATTTTCTGACAGAGAGCTTTGGAAGACTGAGAAACTTTACAGAGAGCCGCA 8875
QY 1288 GAGTGCTCTCTGAGAGAGAGCCAGAAATGTCATCTGGCTCTTCAAGAGAGAGCTGAG 1347
DB 8876 GAGTGCTCTCTGAGAGAGAGCCAGAAATGTCATCTGGCTCTTCAAGAGAGAGCTGAG 8935
QY 1348 GAGTGCTATCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGATGCGCAGAGAAAAATA 1407
DB 8936 GAGTGCTATCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGATGCGCAGAGAAAAATA 8995
QY 1408 GATGAGAGCCTTGAAGAACTCCAGAGAACTTCAAGAGGCGCAGAGAGTGAAGCTGAG 1467
DB 8996 GATGAGAGCCTTGAAGAACTCCAGAGAACTTCAAGAGGCGCAGAGAGTGAAGCTGAG 9055
QY 1468 CTGCGGCAAGTGAAGTGAATCAAGGAGTCTTGGCAAGCCGTTGGGAGTCTCTCATTAAC 1527
DB 9056 CTGCGGCAAGTGAAGTGAATCAAGGAGTCTTGGCAAGCCGTTGGGAGTCTCTCATTAAC 9115
QY 1528 TCTTCTCAAGATCACTCGAAGAACTCAAGGCACTTTCAGAGGAAATTCGCTCTGAAA 1587
DB 9116 TCTTCTCAAGATCACTCGAAGAACTCAAGGCACTTTCAGAGGAAATTCGCTCTGAAA 9175
QY 1588 GAGAACGTGAGCAGCTCAATGACCTTGTCCGCAAGCTTACCACTTTGGGCAATTCAGCTC 1647
DB 9176 GAGAACGTGAGCAGCTCAATGACCTTGTCCGCAAGCTTACCACTTTGGGCAATTCAGCTC 9235
QY 1648 TACCGTATTAACCTCAGCACTCTGAAAGCTGAAACCCAGATGGAACCTTCTGAGGTG 1707
DB 9236 TACCGTATTAACCTCAGCACTCTGAAAGCTGAAACCCAGATGGAACCTTCTGAGGTG 9295
QY 1708 GCGGTGAGAGCAGGTGAGAGGAGCTGATGAAGCCACAGGAACTTTGTGTCCAGCACT 1767
DB 9296 GCGGTGAGAGCAGGTGAGAGGAGCTGATGAAGCCACAGGAACTTTGTGTCCAGCACT 9355
QY 1768 CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCAAAACAA 1827
DB 9356 CAGCACTTTCTTCCACGTCTGTCCAGGGTCCCTGGAGAGAGCCATCTCGCAAAACAA 9415
QY 1828 GTGCGCTACTATATCAACAGAGACTCAAAACAACTGTGTGGGACATCCCAAAATGCA 1887
DB 9416 GTGCGCTACTATATCAACAGAGACTCAAAACAACTGTGTGGGACATCCCAAAATGCA 9475
QY 1888 GAGCTTACAGTCTTTAGCTGACCTGAATATGTCAGATTTTCAGTTATAGACTGCC 1947

DB 9476 GAGCTTACAGCTCTTTAGCTGACCTGAAATATGTCAGATTTTCAGCTTATAGACTGCC 9535
QY 1948 ATGAACTCCGAAAGTGCAGAGAGGCGCTTTGGTTGATCTCTGAGCTGTACAGTGA 2007
DB 9536 ATGAACTCCGAAAGTGCAGAGAGGCGCTTTGGTTGATCTCTGAGCTGTACAGTGA 9595
QY 2008 TGTGATGCTTGGACCAAGCAACCTTCAGACAAAAAGACAGCCCATGATATCTGGAG 2067
DB 9596 TGTGATGCTTGGACCAAGCAACCTTCAGACAAAAAGACAGCCCATGATATCTGGAG 9655
QY 2068 ATTATTAATTTGTTGACCACTATTTATGACCGGCTGAGACCAAGCAACAAATTTGTC 2127
DB 9656 ATTATTAATTTGTTGACCACTATTTATGACCGGCTGAGACCAAGCAACAAATTTGTC 9715
QY 2128 AACGTCCCTCTCTGCGTGAATATGTGTCTGAACTGCTGCTGAATTTATGATAGGGA 2187
DB 9716 AACGTCCCTCTCTGCGTGAATATGTGTCTGAACTGCTGCTGAATTTATGATAGGGA 9775
QY 2188 CGAACAGGAGAGATCCGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA 2247
DB 9776 CGAACAGGAGAGATCCGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA 9835
QY 2248 CATTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGAAGTTCAACAGGATTTGT 2307
DB 9836 CATTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGAAGTTCAACAGGATTTGT 9895
QY 2308 GACCAAGGAGAGCT-GGGCTCTCTGATGATTTTATCCAAATTTCCAAAGACGTTGGG 2366
DB 9896 GACCAAGGAGAGCTGGGCTCTCTGATGATTTTATCCAAATTTCCAAAGACGTTGGG 9955
QY 2367 TGAAGTTGATCCTTTGGGGGAGTAACTTGAAGC 2401
DB 9956 TGAAGTTGATCCTTTGGGGGAGTAACTTGAAGC 9990

RESULT 3
US-08-836-022A-10/c
Sequence 10, Application US/08836022A
Patent No. 6001557

GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
APPLICANT: Wilson, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Chen, Shu-Jen
APPLICANT: Weltzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,022A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVFN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-836-022A-10

Query Match 62.8%; Score 1506.8; DB 3; Length 19307;
-Best Local Similarity 90.6%; Pred. No. 0;
Matches 1607; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

628 TGGGCAACATCTGTAGTGAACAGAGACCGCTGGTCTTTTACAGCACTCATAGA 687
6478 TGGGCAACATCTGTAGTGAACAGAGACCGCTGGTCTTTTACAGCACTCATAGA 6419
668 TTACTGCAAGTTCCTGAGCTGGAGAAAGTTTGTGCTGCTTACAGAAAGTGA 747
6418 TTACTGCAAGTTCCTGAGCTGGAGAAAGTTTGTGCTGCTTACAGAAAGTGA 6359
748 ACAACTGCAAGTTCCTGAGCTGGAGAAAGTTTGTGCTGCTTACAGAAAGTGA 807
6358 ACAACTGCAAGTTCCTGAGCTGGAGAAAGTTTGTGCTGCTTACAGAAAGTGA 6299
808 GAGTAAAGAGCTGATGAACATATGCAAGACCTCCAGTGAATGAAGTCCACCA 867
6298 GAGTAAAGAGCTGATGAACATATGCAAGACCTCCAGTGAATGAAGTCCACCA 6239
868 GATGTTATCAACCTGATGATGAACAGCCAAAGTTCCTGAGTCCCTGAGTCC 927
6238 GATGTTATCAACCTGATGATGAACAGCCAAAGTTCCTGAGTCCCTGAGTCC 6179
928 GATGTTATCAACCTGATGATGAACAGCCAAAGTTCCTGAGTCCCTGAGTCC 987
6178 GATGTTATCAACCTGATGATGAACAGCCAAAGTTCCTGAGTCCCTGAGTCC 6119
988 CGAATAAAGCTCTCAACATTTAGTCCATTTGAGACCGATTCGACAGTGAAGCT 1047
6118 CGAATAAAGCTCTCAACATTTAGTCCATTTGAGACCGATTCGACAGTGAAGCT 6059
1048 CTGCACTTTCTGCAAGAACTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1107
6058 CTGCACTTTCTGCAAGAACTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 5999
1108 CGGAGAGCACTTATGAGGAGCTTCCAGCACTTGAAGAGCAAGATGATGATG 1167
5998 CGTCAAGCACTTATGAGGAGCTTCCAGCACTTGAAGAGCAAGATGATGATG 5939
1168 GCTTCAAGAGGAAATGAAATCTAAGAACTGTAATCAATGATCTCTTGAAGCTGA 1227
5938 GCTTCAAGAGGAAATGAAATCTAAGAACTGTAATCAATGATCTCTTGAAGCTGA 5879
1228 CGAATATTTCTGAGAGAGAGCTTGAAGAGCACTTGAAGAGCACTTGAAGAG 1287
5878 AGAATATTTCTGAGAGAGAGCTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5819
1288 GAGTGCCTCTGAGAGAGAGCTTGAAGAGCACTTGAAGAGCACTTGAAGAG 1347
5818 GAGTGCCTCTGAGAGAGAGCTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5759
1348 GAGTGCCTCTGAGAGAGAGCTTGAAGAGCACTTGAAGAGCACTTGAAGAG 1407
5758 GAGTGCCTCTGAGAGAGAGCTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5699
1408 GATGAGAGCTTGAAGAGAGCTTGAAGAGCACTTGAAGAGCACTTGAAGAG 1467
5698 GATGAGAGCTTGAAGAGAGCTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5639
1468 CTGCGCCAAAGCTGAGGATGATCAAGGATCTGCGAGAGCTGCGAGCTTCTCAT 1527
5638 TTGCGCCAAAGCTGAGGATGATCAAGGATCTGCGAGAGCTGCGAGCTTCTCAT 5579

QY 1528 TCTCTCAAGATCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAG 1587
DB 5578 TCTCTCAAGATCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAG 5519
QY 1588 GAGAGAGTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAG 1647
DB 5518 GAGAGAGTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAG 5459
QY 1648 TCACCTTATTAACCTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 1707
DB 5458 TCACCTTATTAACCTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5399
QY 1708 GCGCTGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 1767
DB 5398 GCGCTGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5339
QY 1768 GAGCACTTCTTCCAGCTCTGCTGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 1827
DB 5338 GAGCACTTCTTCCAGCTCTGCTGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5279
QY 1828 GTGCTCTATATATCAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 1887
DB 5278 GTGCTCTATATATCAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5219
QY 1888 GAGCTCTACAGCTCTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 1947
DB 5218 GAGCTCTACAGCTCTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5159
QY 1948 ATGAAATCTCCAGAGCTGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 2007
DB 5158 ATGAAATCTCCAGAGCTGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5099
QY 2008 TGTGATGCTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 2067
DB 5098 TGTGATGCTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 5039
QY 2068 ATTATTAATGTTGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 2127
DB 5038 ATTATTAATGTTGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 4979
QY 2128 AACGTCCTCTGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 2187
DB 4978 AACGTCCTCTGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 4919
QY 2188 CGAAGAGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 2247
DB 4918 CGAAGAGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 4859
QY 2248 CATTTGAGAGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 2307
DB 4858 CATTTGAGAGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 4799
QY 2308 GACGAGAGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 2367
DB 4798 GACGAGAGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 4739
QY 2368 GAGTGCCTCTGAGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 2401
DB 4738 GAGTGCCTCTGAGAGAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAGCACTTGAAGAG 4705

RESULT 4

US-09-427-048A-10/c
Sequence 10, Application US/09427048A
Patent No. 6203975

GENERAL INFORMATION:

APPLICANT: Trustees of the University of Pennsylvania
Wilson, James M.
Fisher, Krishna J.
Chen, Shu-Jen

TITLE OF INVENTION: Improved Adenovirus Virus and
Methods of Use Thereof

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVEN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10
Query Match 62.8%; Score 1506.8; DB 3; Length 19307;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1607; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 1108 CGGAGGACCTTATGAGGCGACTTCCAGCTTCAGAGGAGACGATGTACATPAG 1167
DB 5998 CGTCAGGACACCCATCGGTGTGATTTCCAGCAGTTCAAGAGCAGATATATACATPAG 5939
QY 1168 GCGTCAAGAGGAGATTTGAATAAGAAACCTGTAATCATAGTACTCTTGAGACTGTA 1227
DB 5938 GCGTCAAGAGGAGATTTGAATAAGAAACCTGTAATCATAGTACTCTTGAGACTGTA 5879
QY 1228 CGAATATTTCTGACAGAGCAGCTTTGAGAGGACTGAGAACTTACAGAGAGCCAG 1287
DB 5878 AGAATATTTCTGACAGAGCAGCTTTGAGAGGACTGAGAACTTACAGAGAGCCAG 5819
QY 1288 GAGTGCCTCTTGAAGAGAGAGCCAGAAATGTACTGCTTGAAGAAAGCGCTGAG 1347
DB 5818 GAATGCCTCTTGAAGAGAGAGCCAGAAATGTACTGCTTGAAGAAAGCGCTGAG 5759
QY 1348 GAGTCAATCTAGAGTGGGAAATTTGAACCTGACTGCTGAGAGAGAAATTA 1407
DB 5758 GAGTCAACGCTGAATGGGAAATTTGAACCTGCTGAGCTGATTTGGAGAGAAATTA 5659
QY 1408 GATGAGACCTTGAAGAGCTCCAGAACTTCAAGAGGCCAGATGAGCTGACCTGAG 1467
DB 5698 GATGAGACCTTGAAGAGCTCCAGAACTTCAAGAGGCCAGATGAGACCTGAG 5639
QY 1468 CTGGGCCAGCTGAGGATCAAGGAGATCTGGCAGCCGCTGGGAGATCTCTCATTTGAC 1527
DB 5638 TTGGGCCAGCTGAGGATCAAGGAGATCTGGCAGCCGCTGGGAGATCTCTCATTTGAC 5579
QY 1528 TCTCTCCAGATCACTTCGAGAAAGTCAAGGCACTTCAAGAGAAATTCGCTGAAA 1587
DB 5578 TCTCTGCAAGATCACTTCGAGAAAGTCAAGGCACTTCAAGAGAAATTCGCTGAAA 5519
QY 1588 GAGAAGTGAAGCCAGCTGATGACCTTGTGCGACCTTACCACTTTGGGAGATTCAGC 1647
DB 5518 GAGAAGTGAAGCCAGCTGATGACCTTGTGCGACCTTACCACTTTGGGAGATTCAGC 5459
QY 1648 TCACCGATATACCTGACGACTCTGAGAGAGCTGAAACCAAGATGAGAGCTTCTGAGGTG 1707
DB 5458 TCACCGATATACCTGACGACTCTGAGAGAGCTGAAACCAAGATGAGAGCTTCTGAGGTG 5399
QY 1708 GCCGTGAGAGACCGAGTCAAGGAGCTGATGAAGCCCAAGGACCTTGGTCCAGATCT 1767
DB 5398 GCCGTGAGAGACCGAGTCAAGGAGCTGATGAAGCCCAAGGACCTTGGTCCAGATCT 5339
QY 1768 CAGACCTTTTCCAGAGCTGTCGACGAGGCTCCCTGAGAGAGAGCCATCTCGCAACAA 1827
DB 5338 CAGACCTTTTCCAGAGCTGTCGACGAGGCTCCCTGAGAGAGAGCCATCTCAACAA 5279
QY 1828 GTGCCCTACTATATCAACGAGAGCTCAACCAACTTGTGAGACCATCCAAATGAC 1887
DB 5278 GTGCCCTACTATATCAACGAGAGCTCAACCAACTTGTGAGACCATCCAAATGAC 5219
QY 1888 GAGCTCTACAGCTTTTACCTGACCTGATATATGATGATTCAGCTTATGAGACTGCG 1947
DB 5218 GAGCTCTACAGCTTTTACCTGACCTGATATATGATGATTCAGCTTATGAGACTGCG 5159
QY 1948 ATGAAATCCGAGAGCTGAGAGAGGCTTTGTGATGATCTTGAAGCTGTCAGCTGCA 2007
DB 5158 ATGAAATCCGAGAGCTGAGAGAGGCTTTGTGATGATCTTGAAGCTGTCAGCTGCA 5099
QY 2008 TGTGATGCTTGAAGAGAGCAACCTCAAGCAAAATGACAGCCCATGATATCTGAG 2067
DB 5098 TGTGATGCTTGAAGAGAGCAACCTCAAGCAAAATGACAGCCCATGATATCTGAG 5039
QY 2068 ATATTAATTTGTTGACCACTATTTATGACCCCTGAGAGAGAGCAAAATTTGATC 2127
DB 5038 ATATTAATTTGTTGACCACTATTTATGATGATGATGATGATGATGATGATGATG 4979
QY 2128 AACGTCCCTCTCTGCTGATATGATGATGATGATGATGATGATGATGATGATGATG 2187
DB 4978 AACGTCCCTCTCTGCTGATATGATGATGATGATGATGATGATGATGATGATGATG 4919
QY 2188 CGAAGAGGAGAGATCGGTGCTGCTTTTAAACTGAGCATATTCCTGATGTAAGCA 2247

Db 4918 CGAAGAGAGAGATCGGTCTCTTTAACTGGCATCATTTCTGTGAAGCA 4859
QY 2248 CATTGAGAGACAATACATACATCTTTCAAGCAAGTGGCAAGTTCAACAGATTTGT 2307
Db 4858 CACTTGAAGACAACTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGT 4799
QY 2308 GACCAAGCGAGGCTGGGCTCTCTTGTGATGATTTCTAATTCACAGAGTTGGT 2367
Db 4798 GACCAAGCGAGGCTGGGCTCTCTTGTGATGATTTCTAATTCACAGAGTTGGT 4739
QY 2368 GAAGTTGATCCTTTGGGGGAGTACATTTGAGC 2401
Db 4738 GAAGTTGATCCTTTGGGGGAGTACATTTGAGC 4705

RESULT 5
US-09-091-501B-7
Sequence 7, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Davies, Kay E
APPLICANT: Tinsley, Jonathan M
TITLE OF INVENTION: Utrrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 6045
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(6037)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
FEATURE:
NAME/KEY: misc feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7

Query Match 28.8%; Score 690.6; DB 4; Length 6045;
Best Local Similarity 60.5%; Pred. No. 1,1e-202;
Matches 1159; Conservative 0; Mismatches 749; Indels 9; Gaps 1;

QY 494 TCGTTCAGAGATCTAGAACAGAGCAAGTCAAGGTCATCTCTCACTCAATGTGG 553
Db 2871 TACTGCTTCAGAGATCTAGAACAGAGTATGATGCTATGCGCTTCATACGCTCTGC 2930
QY 554 TCGTTCAGAGATCTAGAACAGAGTCAAGGTCATCTCTCACTCAATGTGG 613
Db 2931 AGAACTCTCTGAGAGATATGAGAGTATGACACAGAGATGTGAAGAAACACAGAGT 2990
QY 614 TATGGAGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTAC 673
Db 2991 ACTTAAACATCATGATGATCATCTCAACAAAGTATGCTGACAGACAGACGCTTGG 3050
QY 674 AAGACACTATGATTTACTGCAAGAGTTCCCTGACCTGGAAGAGTTCTTGTCTGGC 733
Db 3051 AAGCTGAGTGAAGAGCGTGTGACGCTCTGACAGAGATGTGAAAGCTTCTGAAAGTGA 3110
QY 734 TTACAGAAAGCTGAACAACTGGCAATGTCTACAGAGATGCTACCGGTAAAGAGGCTCC 793

Db 3111 TCCAGAAAGCAAGACCAAGTGAATGTCTGTGAGTCTCTCATGCGAGATGCTC 3170
QY 794 TAGAAGACTCCAAAGGAGTAAAGAGCTGATGAACAAATGGCAGACCTCCAGGTGA 853
Db 3171 TTGAGATGATATCTTGGCCAGGAACTCAACAGACAGATGAGACATCCAGGAGAAA 3230
QY 854 TTGAAGCTCAACAGATGTTTATCAACCTGATGAAAGCCCAAAATCCAGAT 913
Db 3231 TTGATGCCCAATGACATATTTAAAGCATTTGAAGGAAACGAGAGAGATGTTAAAG 3290
QY 914 CCGTGAAGGTTCCGATGATGACATCTCTGTCAAGACGTTGATTAACATGAATTTCA 973
Db 3291 CTTTGGGAATTTCTGAAGAGCTCATATGCTTCAACATGATGATGATGAACAAA 3350
QY 974 AGTGAAGTAACTTGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCAAGTTCTG 1033
Db 3351 GATGGAATACCTTAAAGCAAAATCTGTGATGATGAGGCCCATTTGAGGCGAGCTG 3410
QY 1034 ACCAGTGAAGGCTGTGACACTTCTCTGACAGAACTTCTGTGTGCTACAGCTGAAG 1093
Db 3411 AGAAGTGAACAGTTGCTGATGTCTTGAAGAACTGATCAAAATGCTGAATGAAG 3470
QY 1094 ATGATGAATTAAGCCGAGGAGCACTATTGAGGCGACTTCCAGCACTTCAAGACGA 1153
Db 3471 ATGAAGAGCTTAAAGAAACAATGCTATTGAGAGAGATGTTCCAGCTTACAGCTCAGT 3530
QY 1154 ACATATATAGAGGCTTCAAGAGGAATTTGAACCTGAATCATGATGA 1213
Db 3531 ATACCATTTGAAGGCTTGAAGCGGAGTTAAAGAGAGAAATATCTGTCTGAATG 3590
QY 1214 CTCTTGAAGCTGATCAATATTTCTGACAGACAGCTTT-----GGAAGAGCTAG 1264
Db 3591 CTGTGACAGAGCGGAGTTTCTTGGCTGATCAAGCAATGAGGCCCTGAAAGCCAA 3650
QY 1265 AGAATCTTACAGAGAGCCCAAGAGCTGCTCTGAGAGAGAGCCCAAGATGTCATC 1324
Db 3651 GAAGAACTTACATCAATCAAAAGCAAGATTAATCTGAGAGAGAGCCCAAGATTTGCA 3710
QY 1325 GCGTTTACAGAAAGCAGCTGAGAGGTCATATCTGAGGAGAAATTTGAACCTGCACT 1384
Db 3711 AAGCCATGGGCAACAGTCTTCTGAGTCAAGAAATTTGGAAGTCTAAATGCTGTA 3770
QY 1385 CCGCTGACCTGACAGAGAAATATGATGAGACCTTTGAAGATCTCAGAGATCTCAAG 1444
Db 3771 CTAGCATTTGGCAAGAGCAAGTGAACAGGCTTTGAGAGATCTCAGAGCTGAGGAG 3830
QY 1445 CCAAGATAGAGCTGACCTCAAGCTGCGCAGAGTGAAGTGAAGGATCTGTGAGC 1504
Db 3831 CTATGATGACCTGAGAGCTGACATGAGAGAGAGAGAGTCCGTGCGGAATGCTGAGAGC 3890
QY 1505 CCGTGGAGATCTCTCATTTGACCTCTTCAAGATCACTCGAGAGAAAGTCAAGGACATT 1564
Db 3891 CCGTGGAGACTTACTCATTTGACCTGCTCAGAGATCACTTGAAGAAATCATGAGCATTTA 3950
QY 1565 GAGAGAAATTTGGGCTCTGAAAGAGAAAGTGAAGCAAGTCAATGACCTTGTCTCCAGC 1624
Db 3951 GAGAGAAATTTGAGCAATCACTTTAAAGTTAAAGGTAATGTTTATCAAGTCAAGC 4010
QY 1625 TTACCACTTTGGGATTCAGCTCTCACTCATTAATCACTGACATCTGAGAGAGCTGAACA 1684
Db 4011 TGTCTCACTGAGCTGAGAGTCTCTCAAGATGATGCTGAGAGTGAAGTGAAGTGA 4070
QY 1685 CAGATGAGAGCTTCTGAGAGTGGCGCTGAGAGACCGAGTCAAGGAGCTGAGAGAGCC 1744
Db 4071 TCGATGAGAACTTTTACAGGTTTCTGAGATGATGCTTAAACAGCTTCAAGAGAGCC 4130
QY 1745 ACAGGAGCTTTGCTCAGAGATCTCAGACATTTCTTCAAGCTGTCTGCTCAAGGCTCTGAG 1804
Db 4131 ACAGAGATTTTGAACATCTCTCAGACATTTCTCTCAAGCTGAGTCAAGCTGAGCTGAGC 4190
QY 1805 AGAGAGCATCTTCCCAACAAAGTGCCTTACTATATCAACAAGAGATCAACAACTT 1864

4191 AAGATCCATTTTACATAATAAAGTCCCTTATCATCAACCAACCAACACAGCAACCT 4250
1865 GCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTTACCTGACCTGATATATGCA 1924
4251 GTTGGGACCATCTAAATGACGAACTCTTCAATCCCTGCTGACCTGAAATATGAC 4310
1925 GATCTCAGCTTATAGGACTGCCATGAACTCCGAGACTCGAGAAGCCCTTTCTTGG 1984
4311 GTTTTCTGCTACCTACCTAGCAATCAAAATCCGAAGACTCAAAAAGCACTATCTTGG 4370
1985 ATCTCTTGAGCTGTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2044
4371 ATCTCTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4430
2045 ACCAGCCCATGATATCCCTGCAATTAATTAATTTGTTGACCACTATTATGACCCCTGG 2104
4431 ACCAGCTCTCAGTGTTCAGATGATCACTCACTGCTGACACAACTTATGATGACTTG 4490
2105 AGCAGAGCACAACAATTTGGTCAAGCTCCCTCTCTGCTGATGATGATGATGATGATG 2164
4491 AGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4550
2165 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2224
4551 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4610
2225 GCATCATTTCCCTGTGTAAGCACATTTGGAAGACAGTACAGATACCTTTTCAAGCAAG 2284
4611 GATTAATGCTCTCTCCAAAGCTCTCTGGAAGAAATACAGATATCTCTTTAAGGAAG 4670
2285 TGGCAAGTTCAACAGATTTTGGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2344
4671 TTGCGGGCCGACAGAAATGTTGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4730
2345 TCCAAATTTCAAGACAGTGGGTGAAGTTGCACTCTTTGGGGGACAGTAAACATGAGC 2401
4731 TCCAGATCCCCCGCAGCTAGTGAAGTAGCAGCTTTTGGAGGACAGTAAATTTGAGC 4787

RESULT 6

US-09-091-501B-9
Sequence 9, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 10320
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(10312)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Full length
OTHER INFORMATION: utrophin construct
FEATURE:
NAME/KEY: misc feature
LOCATION: (724)..(758)

OTHER INFORMATION: Precise residue is left open
US-09-091-501B-9

Query Match 28.8%; Score 690.6; DB 4; Length 10320;
Best Local Similarity 60.5%; Pred. No. 1.6e-202;
Matches 1159; Conservative 0; Mismatches 749; Indels 9; Gaps 1;
QY 494 TGCTTCAAGAAGATCTAGAACAAAGTCAAGTCAAGGTCATTTCTCTCACTCAATGCTGG 553
Db 7146 TACTGCTTCAGAACTGGTCTCTGGAGATGGTATCGTCATGGCGTTCGATAAGTCTGC 7205
QY 554 TGTGATGTGATGATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAACAACTTAAGG 613
Db 7206 AGAACTCTCTGGAGGAATATGGAGTGTGACAAAGGATGTGAAGAAACACACAGT 7265
QY 614 TATTGGAGATCGATGGCAAAACATCTGTAGATGGACAGAACCCCTGGTCTTTTAC 673
Db 7266 ACTTAAACATCATGATCAATCTCAACAAAGTATTTGCTGACAGACAGAACCCCTGG 7325
QY 674 AAGACACTCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTTGGCTGGC 733
Db 7326 AGGTGAGTGGAGACGGTGCAGGCTCTCGCAGAGATCTGGAAACTTCTCGAAGTGA 7385
QY 734 TTACAGAGCTGAAACAACTGCCAATGTCTACAGATGCTACCCCTAAGGAAGGCTCC 793
Db 7386 TCAAGAGCAGAGACCAAGTGAATGTGCTGTGATGCTCTCATCGGAGAAATGCTC 7445
QY 794 TAGAAGACTCCAAGGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCCAAGGTGAAA 853
Db 7446 TTCAAGATGATATCTTGGCCAGGAACTCAACAGCAGATGTCAGGACATCCAGGCAAAA 7505
QY 854 TTGAAGCTCACAGATGTTTATCACACTGGATGAAGAACAGCCAAATAATCTCTGAGAT 913
Db 7506 TTGATGCCCAATGACATATTTAAAGCAATTTGACGAAACAGGCAAGAGATGTTAAAG 7565
QY 914 CCCTGGAAGTTCGGATGATGAGTCTCTGTACAAAGAGCTTTGGATAACATGAATTTCA 973
Db 7566 CTTTGGAAATCTGAGAGGCTACTATGCTTCAACATGACTGGATGATATGAACAAA 7625
QY 974 AGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATAGTCCCATTTGGAGCCAGTCTTG 1033
Db 7626 GATGGAATGACTTAAAGCAAAATCTGTAGCATCAGGGCCCATTTTGGAGGCGCAGCTG 7685
QY 1034 ACCAGTGGAGCGTCTGCACCTTTCTCGCAGAACTTCTGCTGTGGCTACAGCTGAAG 1093
Db 7686 AGAAGTGGAAACAGGTTGCTGATGCTTTAGAGAACTGATCAATGGCTGAATATGAAG 7745
QY 1094 ATGATGAATTAAGCCCGCAGGCACTTATTGGAGGCACTTTCCAGCAGTTTCAGAGCAGA 1153
Db 7746 ATGAGAGCTTAAAGAAACAAATGCTTATGGAGGAGATGTTCCAGCCTTACAGCTCCA 7805
QY 1154 ACCATGATAGGCGCTTCAAGAGGAAATGAAGAACTAAAGAACTGTAATCATGAGTA 1213
Db 7806 ATGACCAATTTAAGGCGCTTGAAGCGGAGTTAAAGGAGAAAGAAATATTCTGCTGAAT 7865
QY 1214 CTCTTGAGACTGTACCAATATTCTGACAGAGCAGCTTTT-----GGAAGGACTAG 1264
Db 7866 CTCTGACAGGCGCGAGTTTCTTGGCTGATGACCAATTCAGGCCCCCTGAGAGGCCAA 7925
QY 1265 AGAACTTACAGAGGCGCGAGAGCTGCTCTCTGAGAGAGAGCCAGAAATGTCATC 1324
Db 7926 GAAGAAACCTTACAAATCAAAACAGAAATTAACCTCTGAGGAGAGAGCCCAAGATTGCCA 7985
QY 1325 GGCTTCTAGGAAGCAGGCTGAGGAGTCAATACTGAGTGGGAAATTTGACCTGCTACT 1384
Db 7986 AAGCCATGCGCAACAGTCTTCTGAAGTCAAGAAATTTGGAAAGTCTTAATGCTGTHA 8045
QY 1385 CCCTGACTGGCAGAGAAATAATAGATGAGACCTTGAAGAACTCCAGGAACTTCAAGAGG 1444
Db 8046 CTAGCAATTTGGCAAAAGCAAGTGAAGCAAGGCTTGGAGAACTCAGAGACCTTGCAGGAG 8105
QY 1445 CCAGGATGAGCTGAGCTTCAAGCTGCGCAAGCTGAGTGAATCAAGGATCTCTGCGAGC 1504

Db 8106 CTATGATGACCTGACGCTGACATGAAGGAGGACAGTCCGTGCGAATGCTGGAAGC 8165
QY 1505 CCGTGGGATCTCTCTATTGACTCTCTCAAGATCACTCGAGAAAGTCAAGCACTTC 1564
Db 8166 CCGTGGGAGACTTACTTACTGCTGCTGAGGATCAATGAAAAATCATGSCATTTA 8225
QY 1565 GAGGAGAAATTCGGCTCTGAAAGAGACGAGGAGCCAGCTCAATGACCTTCTCGCCAGC 1624
Db 8226 GAGAGAAATTCAGCAATCAACTTTAAAGTTAAACCGTGAATGATTTATCCAGTCAGC 8285
QY 1625 TTACACATTTGGGCAATTCAGCTCTCACGCTAACTCTCAGCAGCTCTGGAAGCACTGACA 1684
Db 8286 TGTCTCCACTTGACCTGATCCCTCTTAAGATGTCTCGCAGCTAGATGACCTTAATA 8345
QY 1685 CCAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAGGAGCTGCAATGAAGCCC 1744
Db 8346 TCGATGGAAGCTTTTACAGGTTTCTGTGGATGATCGCCTTAAACAGCTTCAGGAAGCCC 8405
QY 1745 ACAGGAGCTTTGGTCCAGCATCTCAGCACTTCTTCCAGCTGTCTCAGGCTCCCTGGG 1804
Db 8406 ACAGAGATTTTGGACCATCTCTCAGCACTTCTCTACGCTCAGTCAGCTGCCGTGGC 8465
QY 1805 ACAGAGCCATCTCGGCAACAAAGTCCCTACTATATCAACACAGAGACTCAAAACAAT 1864
Db 8466 AAGATCCATTTCAATATTAAGTGCCCTATTATCATCAACCATCAACACAGACCCCT 8525
QY 1865 GTGGGACCATCCCAAAATGACAGAGCTCTACGAGCTTTTACGCTGACCTGAAATATGTC 1924
Db 8526 GTTGGACCATCTTAAATGACCGAATCTTTTCAATCCCTTGTGACCTGAATATGTAC 8585
QY 1925 GATCTCAGCTATATAGGACTGCGCAAGTCCGAAAGTCCGAGAGCTCGAGAGGCCCTTTGCTGG 1984
Db 8586 GTTTTCTGCTTACGTCAGCAATCAAAATCCGAGACTACAAAGCACTATGTTTGG 8645
QY 1985 ATCTCTGAGCCTGTGAGCTGATGATGCTTGGACAGCAACACCTCAAGCAAAATG 2044
Db 8646 ATCTCTTAGAGTTGATACAAACAAATGAAATTTTCAACAGCAAAAGTTGAACCAAAATG 8705
QY 2045 ACCAGCCATGATATCTGCGAGATTTAATGTTTGCACCACTATTTATGACCGCTGG 2104
Db 8706 ACCAGCTCTGAGTTTCCAGATGTCATCACTGTCGACCAACTTATGATGAGCTTG 8765
QY 2105 AGCAAGAGCACAAATTTGGTCAACGCTCCCTCTCTGCTGGATATGTCTGAACTGGC 2164
Db 8766 AGCAATGATCAAGGACCTGGTCAAGTCCACTCTGTGTTGATATGTCTCAATGGT 8825
QY 2165 TGTGATGTTTATGATAGCGAGCAAGACGAGGAGATCGGTGCTCTGCTTTTAAACTG 2224
Db 8826 TGTCAATGCTATGACACGGGTGCAACTGGAATAATTAGAGTCAGAGTCTGAAGATTG 8885
QY 2225 GCATCATTTCCCTGTGTAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAG 2284
Db 8886 GATTAATGTTCTCTCCAAAGGTCTCTTGGAGAAAATACAGATATCTCTTAAAGGAG 8945
QY 2285 TGGCAAGTTCAACAGATTTTGTGACCGAGCTGGGCTCCCTTTCTGATGATTTCTA 2344
Db 8946 TTCCGGGGCGGACAGAAATGTGTGACGAGGAGCTGGGCTGTTACTTTCATGATGCCA 9005
QY 2345 TCCAAATTTCCAGACAGTTGGGTGAGTTGATCTCTTTGGGGGAGTAACTGAGC 2401
Db 9006 TCCAGATCCCCGGCAGCTAGGTGAAGTAGCAGCTTTTGGAGGCAATTAATTGAGC 9062

RESULT 7
US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B

; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-091-501B-5

Query Match 3.3%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. 1.3e-14;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 465 CCTAAACCCCAAGTACAAACAATAAGGTGCTTCAAGAAGATCTAGAACAAACAAGT 524
Db 16 CTTGCAAAACCTGCTTGAAGAACATAAAAGTTTGCAAAGTGACCTCGAAGCTGAGCAGT 75
QY 525 CAGGTCATTTCTCTCACTCAGATGCTGGTGTAGTGTATGATCTAGTGGAGATCAGC 584
Db 76 GAAGGTGAATTCCTTAACATCATGCTGTGATGTTGGATGAAAACAGTGGGAGAGCGC 135
QY 585 AACTGCTGCTTTTGGAGAACAACTTAAGGTATTTGGAGATCATGCGCAACATCTGTAG 644
Db 136 CACAGCTGTTTTGGAAGATCAGTTACAGAACTGGGTGAGCGCTGACAGCTGTATGCCG 195
QY 645 ATGGA 649
Db 196 CTGGA 200

RESULT 8
US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-091-501B-4

Query Match 3.3%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 2.3e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 453 TGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAAGGTGCTTCAAGAAGATCTAGA 512
Db 4 TGACCTGCCCTCCCTCGCAGAAGCTGCTTCAAGAACAATAAAAGTTTGCAAAATGACCTTGA 63

QY 513 ACAAGAACAGTCAGGTCATCTCTCTCACTCACTGCTGGTGTAGTGAATCTAG 572
Db 64 AGCTGAACAGGTCGAAGTAAATCTCTTAATCACTCATGCTGGTGTAGTGAATCTAG 123
QY 573 TGGAGATCAGCAGCACTGCTCTTTTGGAGAACCACTTAAAGTATTTGGAGATCGATGGGC 632
Db 124 TGGGAGAGTGCACACAGCTCTTTTGGAGAGATCAGTTACAGAAACTGGTGGCGCTGGAC 183
QY 633 AACATCTGTAGATGGA 649
Db 184 AGCTGTATGCCGCTGGA 200

RESULT 9

US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR FILING DATE: 1995-12-19
; PRIOR FILING DATE: 1995-12-19
; PRIOR FILING DATE: 1996-07-26
; PRIOR FILING DATE: 1996-07-26
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Homo sapiens
JS-09-091-501B-6

Query Match 3.3%; Score 78.6; DB 4; Length 200;

Best Local Similarity 62.4%; Pred. No. 2.3e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 453 TGATCTTGAAGACTAAACGCCAAGTACAAACATAAGGTGCTTCAAGAGATCTAGA 512
Db 4 TGATGTGAATCTCTACAAAGCTGCTAGAGACATAAAAGTTTGCRAAGTATCTGA 63
QY 513 ACAAGAACAGTCAGGTCATCTCTCTCACTCACTGCTGGTGTAGTGAATCTAG 572
Db 64 GGCTGAACAGGTGAAAGTAAATTCATCACTCACTGCTGGTGTAGTGAATCTAG 123
QY 573 TGGAGATCAGCAGCACTGCTCTTTTGGAGAACCACTTAAAGTATTTGGAGATCGATGGGC 632
Db 124 TGGTGAAGCGCTACAGCTATCTCTAGAACACAGTGTACAGAACTGGTGGCGCTGGAC 183
QY 633 AACATCTGTAGATGGA 649
Db 184 AGCAGTATGCCGCTGGA 200

RESULT 10

JS-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFELINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 3.2%; Score 76.6; DB 1; Length 7218;

Best Local Similarity 6.7%; Pred. No. 1.1e-12;
Matches 28; Conservative 236; Mismatches 155; Indels 0; Gaps 0;

QY 113 TTTCTAATGATGTGGAAGTGGTGAAGACCAGGTTTCTATCTCATGAGGGGTACATGATGG 172
Db 1474 TATCTATGCAAGTAGTTAAAGAGATAGAGAATTTGGTACRRRRRRRRRRRRRRRR 1415
QY 173 ATTTCAGAGCCCATCAGGCGGGTGGTGTATTTCTACAAATTGGGAAGTAAGCTGATTG 232
Db 1414 RRR 1355
QY 233 GAACAGGAAATTTATCAGAAAGTGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAA 292
Db 1354 RRR 1295
QY 293 ATTCAAGATGGGAATGCTCAGGTCAGGTAGCTAGTGGAAACAAAGCAATTTACATAGAG 352
Db 1294 RRR 1235
QY 353 TTTTAATGATCTCCAGAACTCAGAAAGTGAAGTGAATGACTGCTAAACAAACAG 412
Db 1234 RRR 1175
QY 413 AAGAAAGAACAGAAATGGAGAGAGCCTTTGAGCCTGATCTTTGAAGACCTAAAC 472
Db 1174 RRR 1115
QY 473 GCCAAGTACAAACATAGTGTCTTCAAGAGATCTAGAACAGCAAGTCAAGGCTC 531
Db 1114 RRR 1056

RESULT 11

US-09-687-875A-13
; Sequence 13, Application US/09687875A

```
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiac, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR FILING DATE: 2000-10-13
; PRIOR FILING DATE: 60/158,868
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pXX-C2 5' junction
US-09-687-875A-13

Query Match          2.8%; Score 63.6; DB 4; Length 238;
Best Local Similarity 94.3%; Pred. No. 1.1e-09;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 924 TTCGGATGATGCACTCTGTACAAAGACGTTTGGATACATGAACCTCAAGTGGAGTGA 983
DB 169 TTCGGACGACGCACTACTGTTCAAAGACGTTTGGATACATGAACCTCAAGTGGAGTGA 228

QY 984 ACTTCGGAAA 993
DB 229 ACTTCGGAAA 238

RESULT 12
US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
US-09-668-313A-10

Query Match          1.8%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 54.7%; Pred. No. 0.0059;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 1801 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 1860
DB 1733 TGGGAGAGAGCTTACACGAGATGGAACTAAGTACTTCAATCAACACGAGACTCAACAGACC 1792

QY 1861 ACTTGCTGGAGACATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCTGAATAT 1920
DB 1793 ACCTCTGGATCCACCCCGCTGATGAGCGCCCTGAACTCTCTCTGTGACAGAGAGTGA 1852

QY 1921 GTCAGATCTTCAGCTTATAGGCTGCCATGAACCTCCGAG 1961
DB 1853 GAGGACTGTCACAGAGAGCTTAACAGACCCGAAAGAGCTGATG 1893

RESULT 13
```

```
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (270001)..(285000)
```

OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (285001)..(300000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (300001)..(315000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (315001)..(330000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t

NAME/KEY: misc feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature

Query Match 1.8%; Score 44; DB 4; Length 1230025;

Best Local Similarity 53.5%; Pred. No. 0.47;

Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 242 AATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGATCTCTTAATTCAAGAT 301
DB 656477 AATTAGAGAAAGAGAGAAAGAGAAATTTGAGGATATCAAGACTCAGATACAAAAT 656418
QY 302 GGGATGCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTTTAATGG 361
DB 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATACGTCCTAGCAGCAATTTATGTGG 656358
QY 362 ATCTCCAGATCAGAAACTGAAGAGTTGATGACTGGCTAACAAAAACAGA 413
DB 656357 CAATTACAGAAAAAACTAAAGCTTCTTAAGAGAGCGCTGGGAAATAGA 656306

RESULT 14
US-09-107-532A-1186
; Sequence 1186, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; City: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Atinello, Pamela Deneké
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1179 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1179
; SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186
Query Match 1.8%; Score 43.4; DB 4; Length 1179;
Best Local Similarity 50.7%; Pred. No. 0.0061;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 371 ATCAGAACTGAAGAGTTGATGCTGCTAACAAAACAGAAAGAAACAGAAAGAA 430
DB 602 AGCAAGATATAAAGATTGATGATCAGACAAAGAAATGAGATACGATCGAGGAA 661
QY 431 TGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTACAAACATA 490
DB 662 TTGTAGAAGTGCTCGTTGGAGGGGTTCCAGCTGGATTAGGAAGCTACGTACAAATGGGACA 721
QY 491 AGTGCTTCAAGAGATCTAGACAGACAACTAGGCTCAATTCCTCTCACTACATGG 550
DB 722 CGAAGCTAGATGCCAAATTCGACAAAGCTGTGGTGTAGTATCAATGCTTTAAAGCGGTAG 781
QY 551 TGGTGGTGTAGTATGATGATCTAGTGG 575
DB 782 AATTGGGTCGGATTCACTTCTGG 806

RESULT 15
US-09-620-312D-69
; Sequence 69, Application US/09620312D
; Patent No. 659662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: Pt_Fl_genes version 1.0
; SEQ ID NO 69
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1522)
US-09-620-312D-69
Query Match 1.8%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.012;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
QY 1428 CCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAGCTGCCCAAGCTGAGTGAT 1487
DB 55 CAAGGAGTTGCACAGGTGGCGGCACACCTGGACGAGCTGGCATGGGTTTCAGGAGCG 114
QY 1488 CAAGGGA---TCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCT 1544
DB 115 GCTGCCACTGGCCATCGACACAGAGCGAGCAACGGTTTCAGGGGGTCCAGCAGCACAT 174
QY 1545 CGAAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAACTGAGGCCACT 1604
DB 175 CAAAAGAACACAGGGCTCGCGCGGAGATCCAGGCGCATGGCGCGCTCGAGGAGGT 234
QY 1605 CAATGACCTTGTCCGACCTTACCACTTTGGGCACTTCACTCTCACCGTATTAACCTCAG 1664
DB 235 GCTGGAGCGCGGGCGCGCTGGCTGCGTCGACGCCCGGAGGAGAGGAGTGGCGCG 294
QY 1665 CACTCTGGAGAGACTGAACACCAAGATGGAAGCTTTTCAGAGTGGCGCTCGAGGACCGAGT 1724
DB 295 GGGCTTGGACACCTGCAGAGCGCTGGCGGAGGCTGGCGGAGGCTGCCGAGCGCGCA 354
QY 1725 CAGCGAGCTGCATGAAGCCCAAGGG 1750
DB 355 GCAGGTGCTGGACCGCGCTTCCAGG 380

Search completed: February 2, 2004, 06:15:54

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 570.303 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-9_COPY_1000_3400

Perfect score: 2401

Sequence: 1 ggcagttcattgtgagag.....tgggggcagtaacattgagc 2401

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2401	100.0	3858	24	Human dystrophin m
2	2401	100.0	4825	24	Adeno-associated v
3	2401	100.0	4848	24	Adeno-associated v
4	2401	100.0	5060	24	Adeno-associated v
5	2250	93.7	3999	24	Human dystrophin m
6	2250	93.7	4966	24	Adeno-associated v
7	2250	93.7	4990	24	Adeno-associated v
8	2067	86.1	4182	24	Human dystrophin m

9	2067	86.1	5149	24	Adeno-associated v
10	1839.4	76.6	5462	24	DNA encoding mini-
11	1737	72.3	3531	24	Human dystrophin m
12	1737	72.3	4498	24	Adeno-associated v
13	1735.6	72.3	5952	22	Human dystrophin g
14	1735.6	72.3	8689	24	DNA encoding mini-
15	1735.6	72.3	11058	24	Human dystrophin p
16	1735.6	72.3	11241	24	cDNA encoding huma
17	1735.6	72.3	11443	24	DNA encoding mini-
18	1735.6	72.3	12923	10	Sequence of human
19	1735.6	72.3	13957	24	Human breast cance
20	1735.6	72.3	13957	24	Human dystrophin g
21	1735.6	72.3	13957	24	Gene #2284 used to
22	1735.6	72.3	13957	24	cDNA encoding huma
23	1725	71.8	5339	24	DNA encoding mini-
24	1724.6	71.8	13977	24	Human bone remodel
25	1723	71.8	2169	24	Human dystrophin r
26	1695	70.6	3510	24	Human dystrophin m
27	1695	70.6	4476	24	Adeno-associated v
28	1506.8	62.8	13815	24	cDNA encoding mous
29	1506.8	62.8	13815	24	Mouse ischaemic co
30	1506.8	62.8	13907	17	Shuttle vector pad
31	1505.2	62.7	13815	19	Mus musculus dyster
32	1419	59.1	5417	24	DNA encoding mini-
33	1376	57.3	1821	24	Human dystrophin r
34	1263	52.6	4414	24	Adeno-associated v
35	1251	52.1	3446	24	Human dystrophin m
36	1058.2	44.1	4402	21	A rod shortened dy
37	989	41.2	1434	24	Human dystrophin r
38	859	35.8	4402	21	A rod shortened dy
39	723.4	30.1	4075	21	A rod shortened dy
40	719.2	30.0	3747	21	A rod shortened dy
41	708.2	29.5	11096	24	cDNA encoding mous
42	692.2	28.8	1991	24	Human dystrophin N
43	690.6	28.8	6045	18	Utrrophin minigene.
44	690.6	28.8	6059	22	Nucleotide sequenc
45	690.6	28.8	10302	24	Human utrophin gen

ALIGNMENTS

RESULT 1
AAD37237
ID AAD37237 standard; DNA; 3858 BP.

AC AAD37237;
XX
DT 21-AUG-2002 (first entry)
XX Human dystrophin minigene delta3849.

DE
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX Homo sapiens.

XX WO200183695-A2.

PD 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -

PS Example 1; Page 48-49; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;

Query Match 100.0%; Score 2401; DB 24; Length 3858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCAGTTTCATGATGAGAGTGAAGTAAACCTGGACCGGTTATCAACACAGCTTTAGAGAA	60
DB	1000	GGCAGTTTCATGATGAGAGTGAAGTAAACCTGGACCGGTTATCAACACAGCTTTAGAGAA	1059
QY	61	GTATTATCGTCTTCTCTGAGGACACATGCGACGACAGGACAGAGATTTCTAAT	120
DB	1060	GTATTATCGTCTTCTCTGAGGACACATGCGACGACAGGACAGAGATTTCTAAT	1119
QY	121	GATGTGGAAGTGTGAAAGACCACTTTCATCTCATGAGGGTACATGATGATTTGACA	180
DB	1120	GATGTGGAAGTGTGAAAGACCACTTTCATCTCATGAGGGTACATGATGATTTGACA	1179
QY	181	GCCCATCAGGCGCGGTGGTAAATATTTACAAATGGGAAGTAAGCTGATGGACAGGA	240
DB	1180	GCCCATCAGGCGCGGTGGTAAATATTTACAAATGGGAAGTAAGCTGATGGACAGGA	1239
QY	241	AAATTATCAGAGATCAAGAACTGAAGTACAAAGACAGATCAATCTCTCAATTAACA	300
DB	1240	AAATTATCAGAGATCAAGAACTGAAGTACAAAGACAGATCAATCTCTCAATTAACA	1299
QY	301	TGGGAATGCTCAGGTAGTCTAGCTAGGAAACAAAGCAATTTACATAGATTTTAATG	360
DB	1300	TGGGAATGCTCAGGTAGTCTAGCTAGGAAACAAAGCAATTTACATAGATTTTAATG	1359
QY	361	GATCTCCAGATCAGAACTGAAAGAGTTGAATGACTGGCTAACAAACACAGAGAAAGA	420
DB	1360	GATCTCCAGATCAGAACTGAAAGAGTTGAATGACTGGCTAACAAACACAGAGAAAGA	1419
QY	421	ACAAGGAAATGGAGGAAGGCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA	480
DB	1420	ACAAGGAAATGGAGGAAGGCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA	1479
QY	481	CACACATAGTGTCTCAAGNAGATCTAGACAAAGACAACTCAGGTCAATTTCTCTC	540
DB	1480	CACACATAGTGTCTCAAGNAGATCTAGACAAAGACAACTCAGGTCAATTTCTCTC	1539
QY	541	ACTCACATGCTGTGTAGTTGATGAATCTAGTGGAGATCAGGCAACTGCTGTTGGAA	600
DB	1540	ACTCACATGCTGTGTAGTTGATGAATCTAGTGGAGATCAGGCAACTGCTGTTGGAA	1599
QY	601	GACACATTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCG	660
DB	1600	GACACATTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCG	1659
QY	661	TGGTCTCTTTTACAGACACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAGA	720
DB	1660	TGGTCTCTTTTACAGACACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAGA	1719

QY	721	TTTCTTGCTGCTTACAGAACTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGT	780
DB	1720	TTTCTTGCTGCTTACAGAACTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGT	1779
QY	781	AAGGAAAGGCTCTAGAGATCCCAAGGAGTAAAGAGCTGATGAAACAATGGCAAGAC	840
DB	1780	AAGGAAAGGCTCTAGAGATCCCAAGGAGTAAAGAGCTGATGAAACAATGGCAAGAC	1839
QY	841	CTCAAGGTCGAATTCGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACACGCCAA	900
DB	1840	CTCAAGGTCGAATTCGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACACGCCAA	1899
QY	901	AAAATCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTGTTACAAAGAGCTTTGGAT	960
DB	1900	AAAATCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTGTTACAAAGAGCTTTGGAT	1959
QY	961	AACATGAATCTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGTGGTCCCATTTG	1020
DB	1960	AACATGAATCTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGTGGTCCCATTTG	2019
QY	1021	GAAGCCAGTTCTGCACAGTGGAGCGCTCTGCACCTTTCTTCAGGAACTTCTGTGTGTG	1080
DB	2020	GAAGCCAGTTCTGCACAGTGGAGCGCTCTGCACCTTTCTTCAGGAACTTCTGTGTGTG	2079
QY	1081	CTACAGCTGAAGATGATGAATTAAGCCGCGCAGGACCTTATTGGAGGCGACTTTCAGCA	1140
DB	2080	CTACAGCTGAAGATGATGAATTAAGCCGCGCAGGACCTTATTGGAGGCGACTTTCAGCA	2139
QY	1141	GTTCAGAGCAGAAACCATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAAGAACCT	1200
DB	2140	GTTCAGAGCAGAAACCATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAAGAACCT	2199
QY	1201	GTAAATCATGAGTACTCTTGAGACTGTACGAATTTTCTGACAGAGCGCTTTGGAGGA	1260
DB	2200	GTAAATCATGAGTACTCTTGAGACTGTACGAATTTTCTGACAGAGCGCTTTGGAGGA	2259
QY	1261	CTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAGAGCCCAAGATGTC	1320
DB	2260	CTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAGAGCCCAAGATGTC	2319
QY	1321	ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTTGAACCTG	1380
DB	2320	ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTTGAACCTG	2379
QY	1381	CATCTCCGCTGACTGGCAGAGAAATAGATGAGCCCTTTGABAGCTCCAGGAACTTCAA	1440
DB	2380	CATCTCCGCTGACTGGCAGAGAAATAGATGAGCCCTTTGABAGCTCCAGGAACTTCAA	2439
QY	1441	GAGGCCACGATGAGTGGACCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTGG	1500
DB	2440	GAGGCCACGATGAGTGGACCTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCCTGG	2499
QY	1501	CAGCCGCTGGGCGATCTCTCATTTGATCTCTCTCAAGATCACTCGAGAACTCAAGGCA	1560
DB	2500	CAGCCGCTGGGCGATCTCTCATTTGATCTCTCTCAAGATCACTCGAGAACTCAAGGCA	2559
QY	1561	CTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGCTCAGCCACGCTCAATGACCTTGTCTG	1620
DB	2560	CTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGCTCAGCCACGCTCAATGACCTTGTCTG	2619
QY	1621	CAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAGACCTG	1680
DB	2620	CAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAGACCTG	2679
QY	1681	AACACACAGATGGAAGCTTTCTGAGGTGGCCGCTCGAGAGCCGAGTCAGGCGATGCAATGAA	1740
DB	2680	AACACACAGATGGAAGCTTTCTGAGGTGGCCGCTCGAGAGCCGAGTCAGGCGATGCAATGAA	2739
QY	1741	GCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCAGCTGTGTCCAGGTCCC	1800
DB	2740	GCCCAACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCAGCTGTGTCCAGGTCCC	2799

QY 661 TGGGTTCTTTTACAGACACTATAGATTACTGCAACAGTTCCCGCTGGACCTCGAAG 720
DB |||||
QY 2417 TGGGTTCTTTTACAGACACTATAGATTACTGCAACAGTTCCCGCTGGACCTCGAAG 2476
DB |||||
QY 721 TTTCTTGCTGGCTTACAGAGCTGAAACAACTGCAATGTCCTACAGGATGCTACCGT 780
DB |||||
QY 2477 TTTCTTGCTGGCTTACAGAGCTGAAACAACTGCAATGTCCTACAGGATGCTACCGT 2536
DB |||||
QY 781 AAGGAAAGGCTCTAGAGACTCCAGAGGATTAAGAGCTGATGAACANTGSCAGAC 840
DB |||||
QY 2537 AAGGAAAGGCTCTAGAGACTCCAGAGGATTAAGAGCTGATGAACAAATGSCAGAC 2596
DB |||||
QY 841 CTCGAAGTGAAATGAAAGCTCACAGAGTGTATCAACACCTGGATGAAACAGCAA 900
DB |||||
QY 2597 CTCGAAGTGAAATGAAAGCTCACAGAGTGTATCAACACCTGGATGAAACAGCAA 2656
DB |||||
QY 901 AAAATCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCCTGTACAAAGAGCTTGGAT 960
DB |||||
QY 2657 AAAATCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCCTGTACAAAGAGCTTGGAT 2716
DB |||||
QY 961 AACATGAATCTCAAGTGGAGTGAACCTTCGGAAGAGTCTCTCAACATTAGTCCCATTTG 1020
DB |||||
QY 2717 AACATGAATCTCAAGTGGAGTGAACCTTCGGAAGAGTCTCTCAACATTAGTCCCATTTG 2776
DB |||||
QY 1021 GAAGCAGTCTGACAGTGGAGCGTCTGCAACCTTCTGCAAGAACTTCTGCTGTGG 1080
DB |||||
QY 2777 GAAGCAGTCTGACAGTGGAGCGTCTGCAACCTTCTGCAAGAACTTCTGCTGTGG 2836
DB |||||
QY 1081 CTACAGCTGAAGATGATGAATTAAGCGGAGGACCTATTGAGGCGACTTCCAGCA 1140
DB |||||
QY 2837 CTACAGCTGAAGATGATGAATTAAGCGGAGGACCTATTGAGGCGACTTCCAGCA 2896
DB |||||
QY 1141 GTTCAGAGCAACGATGTACATAGGCGCTTCAAGAGGGAATTAAGAACTTAAGAACT 1200
DB |||||
QY 2897 GTTCAGAGCAACGATGTACATAGGCGCTTCAAGAGGGAATTAAGAACTTAAGAACT 2956
DB |||||
QY 1201 GTAATCATGACTCTTGAGACTGTAGATATTTCTGACAGAGCAGCTTGGAGGA 1260
DB |||||
QY 2957 GTAATCATGACTCTTGAGACTGTAGATATTTCTGACAGAGCAGCTTGGAGGA 3016
DB |||||
QY 1261 CTAGAGAACTCTACAGAGGCCAGAGAGTGCCTCTGAGGAGAGAGCCAGAAATGTC 1320
DB |||||
QY 3017 CTAGAGAACTCTACAGAGGCCAGAGAGTGCCTCTGAGGAGAGAGCCAGAAATGTC 3076
DB |||||
QY 1321 ACTGGCTTTACGAAAGCAGCTGAGAGGTCAATATCTGAGTGGGAAATTAAGAACTG 1380
DB |||||
QY 3077 ACTGGCTTTACGAAAGCAGCTGAGAGGTCAATATCTGAGTGGGAAATTAAGAACTG 3136
DB |||||
QY 1381 CACTCCGCTGACTGGCAGAGAAATATAGATGAGACCTTTGAAAGACTCCAGAACTTCAA 1440
DB |||||
QY 3137 CACTCCGCTGACTGGCAGAGAAATATAGATGAGACCTTTGAAAGACTCCAGAACTTCAA 3196
DB |||||
QY 1441 GAGGCCAGGATGAGCTGAGCTCAAGCTGCGCCAGAGTGAAGTGAAGGATTCCTGG 1500
DB |||||
QY 3197 GAGGCCAGGATGAGCTGAGCTCAAGCTGCGCCAGAGTGAAGTGAAGGATTCCTGG 3256
DB |||||
QY 1501 CAGCCGCTGGGAGTCTCTCACTGACTCTCCAGATCACTCCAGAAAGTCAAGGCA 1560
DB |||||
QY 3257 CAGCCGCTGGGAGTCTCTCACTGACTCTCCAGATCACTCCAGAAAGTCAAGGCA 3316
DB |||||
QY 1561 CTTTCAGAGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGCTGC 1620
DB |||||
QY 3317 CTTTCAGAGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGCTGC 3376
DB |||||
QY 1621 CAGCTTACCACTTTGGGAGTTCAGCTCTCACGATTAACCTCAGCACTCTGAGAGCTG 1680
DB |||||
QY 3377 CAGCTTACCACTTTGGGAGTTCAGCTCTCACGATTAACCTCAGCACTCTGAGAGCTG 3436
DB |||||
QY 1681 AACACAGATGAAAGCTTCTGAGGTGGCGCTGAGGAGCCGAGTCAGGAGCTGCATGAA 1740
DB |||||
QY 3437 AACACAGATGAAAGCTTCTGAGGTGGCGCTGAGGAGCCGAGTCAGGAGCTGCATGAA 3496
DB |||||
QY 1741 GCGCCAGGAGCTTTGGTCAGCAATCTCAGCACTTTCTTCCACGCTCTGTCAGGAGTCCC 1800
DB |||||

DB 3497 GCCCAGAGGACTTTGGTCCAGCATCTCAGCACTTCTTCCACGCTGTCCAGGTCCTC 3556
QY |||||
DB 1801 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACAGAGACTCAACA 1860
QY |||||
DB 3557 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACAGAGACTCAACA 3616
QY |||||
DB 1861 ACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATAT 1920
QY |||||
DB 3617 ACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATAT 3676
QY |||||
DB 1921 GTCAGATTCTCAGCTTATAGGACTGCGCATGAACTTCGGAAGACTGCGAAGGCCCTTTGC 1980
QY |||||
DB 3677 GTCAGATTCTCAGCTTATAGGACTGCGCATGAACTTCGGAAGACTGCGAAGGCCCTTTGC 3736
QY |||||
DB 1981 TTGGATCTCTTGGGACCTGTCAGTGCATGATGCTTGGACCAACAACCTCAAGCAA 2040
QY |||||
DB 3737 TTGGATCTCTTGGGACCTGTCAGTGCATGATGCTTGGACCAACAACCTCAAGCAA 3796
QY |||||
DB 2041 AATGACAGCCCATGGATATCTCAGATTTAATTTGTTGACCACTATTTATGACCGC 2100
QY |||||
DB 3797 AATGACAGCCCATGGATATCTCAGATTTAATTTGTTGACCACTATTTATGACCGC 3856
QY |||||
DB 2101 CTGGAGCAAGAGCAACAATTTGCTCAAGCTCCCTCTCGTGGATATGCTGAC 2160
QY |||||
DB 3857 CTGGAGCAAGAGCAACAATTTGCTCAAGCTCCCTCTCGTGGATATGCTGAC 3916
QY |||||
DB 2161 TGGCTCTCAATTTTATGATACGGGACGAAACAGGAGGATCCGTGCTCTCTTTTAA 2220
QY |||||
DB 3917 TGGCTCTCAATTTTATGATACGGGACGAAACAGGAGGATCCGTGCTCTCTTTTAA 3976
QY |||||
DB 2221 ACTGGATCATTTCCCTGTTAAAGCAATTTGGAAGCAAGTACAGTACCTTTTCAAG 2280
QY |||||
DB 3977 ACTGGATCATTTCCCTGTTAAAGCAATTTGGAAGCAAGTACAGTACCTTTTCAAG 4036
QY |||||
DB 2281 CAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCAGGCTGGGCTCTCTCTGATGAT 2340
QY |||||
DB 4037 CAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCAGGCTGGGCTCTCTCTGATGAT 4096
QY |||||
DB 2341 TCTATCCAAATTCGAAGACAGTTGGTGAAGTTCATCTTTGGGGGAGTAACTTGAAG 2400
QY |||||
DB 4097 TCTATCCAAATTCGAAGACAGTTGGTGAAGTTCATCTTTGGGGGAGTAACTTGAAG 4156
QY |||||
DB 2401 C 2401
DB 4157 C 4157

RESULT 3

AAD37263
ID AAD37263 standard; DNA; 4848 BP.

XX AAD37263;

XX 21-AUG-2002 (first entry)

XX Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

XX Chimeric - Homo sapiens.

XX Chimeric - Cytomegalovirus.

XX Chimeric - Unidentified.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-20077P.

XX (XIAO/) XIAO X.
XX PA
XX PI
XX Xiao X;
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -
XX
XX Example 1; Page 68-70; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene is operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid construct containing human dystrophin minigenes, a
XX cytomegalovirus (CMV) promoter and a small polyA signal sequence.
XX
XX Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 2401; DB 24; Length 4848;
XX Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 1 GCGAGTTCATGATGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
b 1780 GCGAGTTCATGATGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1839
Y 61 GTATTATCGTGGCTTTCTCTGCTGAGACACATTCGACGACAGGAGAGATTTCTAAT 120
b 1840 GTATTATCGTGGCTTTCTCTGCTGAGACACATTCGACGACAGGAGAGATTTCTAAT 1899
Y 121 GATGTGGAAGTGGTGAAGACAGCTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
b 1900 GATGTGGAAGTGGTGAAGACAGCTTCATCTCATGAGGGGTACATGATGGATTTGACA 1959
Y 181 GCCCATCAGGCGCGGGTGGTAAATTTCTACATTTGGGAAGTAACTGATTTGGAACAGGA 240
b 1960 GCCCATCAGGCGCGGGTGGTAAATTTCTACATTTGGGAAGTAACTGATTTGGAACAGGA 2019
Y 241 AAATTATCAGAGAGTGAAGAACTCAAGTACAGAGCAGATGATCTCTAAATTCAGA 300
b 2020 AAATTATCAGAGAGTGAAGAACTCAAGTACAGAGCAGATGATCTCTAAATTCAGA 2079
Y 301 TGGGAATGCTCAGGGTAGCTAGCATGATGAAACAAAGCAATTTACATPAGAGTTTAAATG 360
b 2080 TGGGAATGCTCAGGGTAGCTAGCATGATGAAACAAAGCAATTTACATPAGAGTTTAAATG 2139
Y 361 GATCTCCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAACACAGAGAAGA 420
b 2140 GATCTCCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAACACAGAGAAGA 2199
Y 421 ACAAGGAAATGGAGAGAGGCTCTTGACCTGATCTTTGAAGACCTTAAACGGCAAGTA 480
b 2200 ACAAGGAAATGGAGAGAGGCTCTTGACCTGATCTTTGAAGACCTTAAACGGCAAGTA 2259
Y 481 CAACAAACATPAGGTGCTCAAGAGATCTAGAACAAAGCAAGTCAGGTCAATTTCTCTC 540
b 2260 CAACAAACATPAGGTGCTCAAGAGATCTAGAACAAAGCAAGTCAGGTCAATTTCTCTC 2319
Y 541 ACTCATAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
b 2320 ACTCATAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2379

QY 601 GAACAACTTAAGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGC 660
DB 2380 GAACAACTTAAGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAAAGCCGC 2439
QY 661 TGGGTCTTTTACAAGACACTCATGATTACTGTCAACAGTTCCTCCCTGGACCTGGAAAAG 720
DB 2440 TGGGTCTTTTACAAGACACTCATGATTACTGTCAACAGTTCCTCCCTGGACCTGGAAAAG 2499
QY 721 TTTCTTGGCTGGCTTACAGAAGCTGAAACAACTGCAATGTCTCTCAGGATGCTACCCGT 780
DB 2500 TTTCTTGGCTGGCTTACAGAAGCTGAAACAACTGCAATGTCTCTCAGGATGCTACCCGT 2559
QY 781 AAGGAAAGGCTCTCTAGAGACTCAAGAGGAGTAAAGAGCTGATGAAACAAATGGCGAAGAC 840
DB 2560 AAGGAAAGGCTCTCTAGAGACTCAAGAGGAGTAAAGAGCTGATGAAACAAATGGCGAAGAC 2619
QY 841 CTCCAAGGTGAATTCAGAGCTCACAGAGTGTTCACAACTGATGAAACAAACAGCCAA 900
DB 2620 CTCCAAGGTGAATTCAGAGCTCACAGAGTGTTCACAACTGATGAAACAAACAGCCAA 2679
QY 901 AAAATCCTGAGATCCCTGGAAGTTCGATGATGCAAGTCTCTTTACAAAGACGTTTGGAT 960
DB 2680 AAAATCCTGAGATCCCTGGAAGTTCGATGATGCAAGTCTCTTTACAAAGACGTTTGGAT 2739
QY 961 AACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCAATTTG 1020
DB 2740 AACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCAATTTG 2799
QY 1021 GAAGCCAGTTCGACCAAGTGGAGGCTCTGCAACCTTTCTCTGAGGAACTTTCTGGTGTGG 1080
DB 2800 GAAGCCAGTTCGACCAAGTGGAGGCTCTGCAACCTTTCTCTGAGGAACTTTCTGGTGTGG 2859
QY 1081 CTACAGCTGAAAGATGATGAAATTAAGCCGGCAGGACCTTATTTGGAGGCGACTTTCCAGCA 1140
DB 2860 CTACAGCTGAAAGATGATGAAATTAAGCCGGCAGGACCTTATTTGGAGGCGACTTTCCAGCA 2919
QY 1141 GTTCAGAAGCAGAACCATGTACATAGGGCTTTCAAGAGGGAATTTGAAACTAAAGAACCT 1200
DB 2920 GTTCAGAAGCAGAACCATGTACATAGGGCTTTCAAGAGGGAATTTGAAACTAAAGAACCT 2979
QY 1201 GTAAATCATGAGTACTCTTGAGACTGTACGAATTTTCTGACAGAGCAGCCTTTGGAAGGA 1260
DB 2980 GTAAATCATGAGTACTCTTGAGACTGTACGAATTTTCTGACAGAGCAGCCTTTGGAAGGA 3039
QY 1261 CTAGAGAACTCTTACCAGAGCCGAGAGGCTGCTCTCTGAGGAGAGAGCCAGAGATGTC 1320
DB 3040 CTAGAGAACTCTTACCAGAGCCGAGAGGCTGCTCTCTGAGGAGAGAGCCAGAGATGTC 3099
QY 1321 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATPACTGAGTGGGAAAAATTTGAACCTG 1380
DB 3100 ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATPACTGAGTGGGAAAAATTTGAACCTG 3159
QY 1381 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAA 1440
DB 3160 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAA 3219
QY 1441 GAGGCCACGATGAGCTGAGCTCAAGCTGCGCCAGGCTGATCAAGGGATCTCTGG 1500
DB 3220 GAGGCCACGATGAGCTGAGCTCAAGCTGCGCCAGGCTGATCAAGGGATCTCTGG 3279
QY 1501 CAGCCCGTGGGCGATCTCTCTCAATGACTCTCTCAGAGTCACTCTGAGAAAGTCAAGGCA 1560
DB 3280 CAGCCCGTGGGCGATCTCTCTCAATGACTCTCTCAGAGTCACTCTGAGAAAGTCAAGGCA 3339
QY 1561 CTTGAGGAGGAAATTTGGGCTCTGAAGAGAGAGTGGCCAGCTCAATGACCTTGTCTGC 1620
DB 3340 CTTGAGGAGGAAATTTGGGCTCTGAAGAGAGAGTGGCCAGCTCAATGACCTTGTCTGC 3399
QY 1621 CAGCTTACCACTTTTGGGCTTCAAGTCTCTCAGCGTAACTCTCAGCACTCTGGAAGACCTG 1680
DB 3400 CAGCTTACCACTTTTGGGCTTCAAGTCTCTCAGCGTAACTCTCAGCACTCTGGAAGACCTG 3459
QY 1681 AACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGGAGCCGAGTCAAGGAGCTGATGAA 1740

DB 3460 AACACACAGATGGAAGCTTCTCAGGTGCGCTCGAGGACCGAGTCAGCGAGCTGCATGAA 3519
QY 1741 GCCACAGGGACTTTGGTCCAGCATCTCAGACATTTCTTCCACGCTGTGTCAGGGTCCC 1800
DB 3520 GCCACAGGGACTTTGGTCCAGCATCTCAGACATTTCTTCCACGCTGTGTCAGGGTCCC 3579
QY 1801 TGGGAGAGGCGCATCTGCCAAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA 1860
DB 3580 TGGGAGAGGCGCATCTGCCAAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA 3639
QY 1861 ACTTCTGGGACCATCCCAAATGACAGAGTCTACAGTCTTTAGCTGAGCTGAATAAT 1920
DB 3640 ACTTCTGGGACCATCCCAAATGACAGAGTCTACAGTCTTTAGCTGAGCTGAATAAT 3699
QY 1921 GTCAGATCTCAGCTTATAGACTGCGCATGAATCCGAGACTGCAGAGGCCCTTTGC 1980
DB 3700 GTCAGATCTCAGCTTATAGACTGCGCATGAATCCGAGACTGCAGAGGCCCTTTGC 3759
QY 1981 TTGGATCTCTTGGAGCTGTGACGCTGTCATGTGATGCTTTGGACCAACACCTCAAGCAA 2040
DB 3760 TTGGATCTCTTGGAGCTGTGACGCTGTCATGTGATGCTTTGGACCAACACCTCAAGCAA 3819
QY 2041 AATGACGACCCCATGATCTCTGAGATTTATTTGACCACTATTATGACGCG 2100
DB 3820 AATGACGACCCCATGATCTCTGAGATTTATTTGACCACTATTATGACGCG 3879
QY 2101 CTGGAGCAAGAGACCAACAATTTGGTCAACGCTCTCTCTGCTGGATATGTGCTGAAC 2160
DB 3880 CTGGAGCAAGAGACCAACAATTTGGTCAACGCTCTCTCTGCTGGATATGTGCTGAAC 3939
QY 2161 TGGCTGCTGAATGTTTATGATACGGGAGCAACAGGGAGATCCGTCCTGCTCTTTAAA 2220
DB 3940 TGGCTGCTGAATGTTTATGATACGGGAGCAACAGGGAGATCCGTCCTGCTCTTTAAA 3999
QY 2221 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTACAGATCTTTTCAAG 2280
DB 4000 ACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTACAGATCTTTTCAAG 4059
QY 2281 CAAAGTGGCAAGTTCAACAGGATTTGTGACGAGCGAGGCTGGGCTCTCTGATGAT 2340
DB 4060 CAAAGTGGCAAGTTCAACAGGATTTGTGACGAGCGAGGCTGGGCTCTCTGATGAT 4119
QY 2341 TCTATCCAAATTCACAGAGTGGGTGAGTGCATCTTTGGGGCAGTAACATTTGAG 2400
DB 4120 TCTATCCAAATTCACAGAGTGGGTGAGTGCATCTTTGGGGCAGTAACATTTGAG 4179
QY 2401 C 2401
DB 4180 C 4180
RESULT 4
AD37264
ID AD37264 standard; DNA; 5060 BP.
XX
AC
AD37264;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.
PF
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
WI; 2002-049342/06.
XX
New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
Example 1; Page 70-71; 71pp; English.
XX
The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter
CC and a small polyA signal sequence.
XX
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;
Query Match 100.0%; Score 2401; DB 24; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGTTCATTGATGAGAGTGAAGTAACTGAACCTGGACCGCTTATCAACAGCTTTTGAAGAA 60
DB 1992 GGCAGTTCATTGATGAGAGTGAAGTAACTGAACCTGGACCGCTTATCAACAGCTTTGAAGAA 2051
QY 61 GTATTATCTGCTGGCTTCTTCTCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT 120
DB 2052 GTATTATCTGCTGGCTTCTTCTCTGAGGACACATTCGAAGCACAGGAGAGATTTCTAAT 2111
QY 121 GATGTGGAAGTGTGAAACACAGCTTCTATCTATGATGAGGGGTACATGATGGATTTGACA 180
DB 2112 GATGTGGAAGTGTGAAACACAGCTTCTATCTATGATGAGGGGTACATGATGGATTTGACA 2171
QY 181 GCCATCAGGCGGGTGTGCTAATATTTCTCAATTTGGGAAGTAAGCTGATTGGAAACAGGA 240
DB 2172 GCCATCAGGCGGGTGTGCTAATATTTCTCAATTTGGGAAGTAAGCTGATTGGAAACAGGA 2231
QY 241 AATTATCAAGATGAAGAACTGAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 300
DB 2232 AATTATCAAGATGAAGAACTGAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2291
QY 301 TGGGAATGCTCAGGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 360
DB 2292 TGGGAATGCTCAGGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 2351
QY 361 GATCTCCAGAACTCAGAACTGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
DB 2352 GATCTCCAGAACTCAGAACTGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2411
QY 421 ACAAGGAATGAGGAGAGGCTCTTGGACCTGATCTTTGAAGACTTAAACGCCAAGTA 480
DB 2412 ACAAGGAATGAGGAGAGGCTCTTGGACCTGATCTTTGAAGACTTAAACGCCAAGTA 2471
QY 481 CAACAAACATAAGGTGCTTCAAGAAAGATCTAGAACAGAAACAAGTCAGGCTCAATTTCTCTC 540
DB 2472 CAACAAACATAAGGTGCTTCAAGAAAGATCTAGAACAGAAACAAGTCAGGCTCAATTTCTCTC 2531

QY 541 ACTCAGATGGTGGTGTAGTGTAGTGAATCTAGTGGAGATCACGAACTGCTGTTTGGAA 600
Db 2532 ACTCAGATGGTGGTGTAGTGTAGTGAATCTAGTGGAGATCACGAACTGCTGTTTGGAA 2591
QY 601 GAACAACTTAAGGTATTTGGAGATCGATGGGAAAACATCTGTAGTGGACAGAGACCOC 660
Db 2592 GAACAACTTAAGGTATTTGGAGATCGATGGGAAAACATCTGTAGTGGACAGAGACCOC 2651
QY 661 TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAG 720
Db 2652 TGGGTTCTTTTACAAGACACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAAG 2711
QY 721 TTTCTTGCCTTACAGAACTGAAAGCTGAAACAACTGCAATGTCTCTACAGATGCTACCCGT 780
Db 2712 TTTCTTGCCTTACAGAACTGAAAGCTGAAACAACTGCAATGTCTCTACAGATGCTACCCGT 2771
QY 781 AAGGAAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGAC 840
Db 2772 AAGGAAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGAC 2831
QY 841 CTCGAGGTGAAATTTGAAGCTCACACAGATGTTTATCAAACTGGATGAAACAGCCAA 900
Db 2832 CTCGAGGTGAAATTTGAAGCTCACACAGATGTTTATCAAACTGGATGAAACAGCCAA 2891
QY 901 AAAATCCTGAGATCCCTGGAGGTTCCGATGATGCACTCTGTCTCAAAAGAGCTTTGGAT 960
Db 2892 AAAATCCTGAGATCCCTGGAGGTTCCGATGATGCACTCTGTCTCAAAAGAGCTTTGGAT 2951
QY 961 AACATGAATTTCAAGTGGAGTGAATTTGGAAAAGTCTCTCAAACTAGGTCCCAATTTG 1020
Db 2952 AACATGAATTTCAAGTGGAGTGAATTTGGAAAAGTCTCTCAAACTAGGTCCCAATTTG 3011
QY 1021 GAAGCCAGTTCTGACCCAGTGAAGAGCTGTGCACCTTTCTGACAGAACTTCTGTGTGG 1080
Db 3012 GAAGCCAGTTCTGACCCAGTGAAGAGCTGTGCACCTTTCTGACAGAACTTCTGTGTGG 3071
QY 1081 CTACAGCTGAAGATGATGAATTAAGCCGAGGACCTATTGGAGCGACTTTCAGCA 1140
Db 3072 CTACAGCTGAAGATGATGAATTAAGCCGAGGACCTATTGGAGCGACTTTCAGCA 3131
QY 1141 GTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAACTAAGAACCT 1200
Db 3132 GTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAACTAAGAACCT 3191
QY 1201 GTAATCATGATGACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGA 1260
Db 3192 GTAATCATGATGACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGA 3251
QY 1261 CTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTCTGAGAGAGAGCCAGAAATGTC 1320
Db 3252 CTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTCTGAGAGAGAGCCAGAAATGTC 3311
QY 1321 ACTCGGCTTCTAGAAAGCAGGCTGAGAGGTCAATCTGAGTGGGAAAAATTTGAACCTG 1380
Db 3312 ACTCGGCTTCTAGAAAGCAGGCTGAGAGGTCAATCTGAGTGGGAAAAATTTGAACCTG 3371
QY 1381 CACTCCGCTGACTGGCAGAGAAAAATAGATGACACCTTTGAAAGACTCCAGGAACCTTCAA 1440
Db 3372 CACTCCGCTGACTGGCAGAGAAAAATAGATGACACCTTTGAAAGACTCCAGGAACCTTCAA 3431
QY 1441 GAGGCCACGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGG 1500
Db 3432 GAGGCCACGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATCCTGG 3491
QY 1501 CAGCCGTGGGCGATCTCTCATGTACTCTCCAGATCATCTCGAGAACTCGAGAACTCAGGCA 1560
Db 3492 CAGCCGTGGGCGATCTCTCATGTACTCTCCAGATCATCTCGAGAACTCGAGAACTCAGGCA 3551
QY 1561 CTTTCGAGGAGAAATTTGGCCTCTCTGAAAGAGAAAGCTGAGGCCAGCTCAATGACCTTGTGCG 1620
Db 3552 CTTTCGAGGAGAAATTTGGCCTCTCTGAAAGAGAAAGCTGAGGCCAGCTCAATGACCTTGTGCG 3611

QY 1621 CAGCTTACCACCTTTGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTG 1680
Db 3612 CAGCTTACCACCTTTGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTG 3671
QY 1681 AACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAA 1740
Db 3672 AACACCAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAA 3731
QY 1741 GCCACACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGCCAGGTCCTC 1800
Db 3732 GCCACACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGCCAGGTCCTC 3791
QY 1801 TGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 1860
Db 3792 TGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 3851
QY 1861 ACTTGTGGGACCAATCCAAAATGACAGAGCTTACCAGTCTTTTAGCTGACCTCAATAAT 1920
Db 3852 ACTTGTGGGACCAATCCAAAATGACAGAGCTTACCAGTCTTTTAGCTGACCTCAATAAT 3911
QY 1921 CTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACTGCAAGAGGCCCTTTGC 1980
Db 3912 CTCAGATTCTCAGCTTATAGACTGCCATGAACTCCGAAGACTGCAAGAGGCCCTTTGC 3971
QY 1981 TTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACGAGCACAACCTCAAGCAA 2040
Db 3972 TTGGATCTCTTGAGCCCTGTCAGCTGCATGTGATGCTTGGACGAGCACAACCTCAAGCAA 4031
QY 2041 AATGACGAGCCCATGGATATCTCTCAGATTATTAATTTGTTGACCACTATTTATGACCGC 2100
Db 4032 AATGACGAGCCCATGGATATCTCTCAGATTATTAATTTGTTGACCACTATTTATGACCGC 4091
QY 2101 CTGAGCAGAGAGCACAACATTTGGTCAAGTCCCTCTCTGCGTGGATATGTCTGAAC 2160
Db 4092 CTGAGCAGAGAGCACAACATTTGGTCAAGTCCCTCTCTGCGTGGATATGTCTGAAC 4151
QY 2161 TGGCTGCTGAATGTTTATGATACGGGACGAACGAGGAGATCCGTGCTGCTTTTAAA 2220
Db 4152 TGGCTGCTGAATGTTTATGATACGGGACGAACGAGGAGATCCGTGCTGCTTTTAAA 4211
QY 2221 ACTGGCATCATTTCCCTGTGTAAAGCATTGGAAGACAAAGTACAGATACCTTTTCAAG 2280
Db 4212 ACTGGCATCATTTCCCTGTGTAAAGCATTGGAAGACAAAGTACAGATACCTTTTCAAG 4271
QY 2281 CAAGTGCAGAGTTCAACAGGATTTGTGACCGCAGGCTGGGCTCCTTTCTGCAATGAT 2340
Db 4272 CAAGTGCAGAGTTCAACAGGATTTGTGACCGCAGGCTGGGCTCCTTTCTGCAATGAT 4331
QY 2341 TCTATCCAAATTCAGACAGATTGGGTGAAGTTGCATTCCTTTGGGGCAGTAACATTGAG 2400
Db 4332 TCTATCCAAATTCAGACAGATTGGGTGAAGTTGCATTCCTTTGGGGCAGTAACATTGAG 4391
QY 2401 C 2401
Db 4392 C 4392
RESULT 5
AAD37234
ID AAD37234 standard; DNA; 3999 BP.
AC AAD37234;
XX
DT 21-AUG-2002 (first entry)
XX Human dystrophin minigene delta3990.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX

Db 2560 GAAAGACTCCAGAACTTCAAGAGGCCAGGATGAGTGGAGCTCAAGCTGGCCAGCT 2619
2y 1480 GAGGTGATCAAGGGATCTCGGAGCCCGTGGGCGATCTCTCATTTGACTCTCTCCAGAT 1539
2b 2620 GAGGTGATCAAGGGATCTCGGAGCCCGTGGGCGATCTCTCATTTGACTCTCTCCAGAT 2679
2y 1540 CACCTCGAGAAAGTCAAGGACACTTCGAGAGAGAAATTCGCGCTCTGAAGAGAGACGTGAGC 1599
2b 2680 CACCTCGAGAAAGTCAAGGACACTTCGAGAGAGAAATTCGCGCTCTGAAGAGAGACGTGAGC 2739
2y 1600 CAGCTCAATGACCTTCTCGCCAGCTTACACATTTGGGCAATTCAGCTCTCACGTATAAC 1659
Db 2740 CAGCTCAATGACCTTCTCGCCAGCTTACACATTTGGGCAATTCAGCTCTCACGTATAAC 2799
2y 1660 CTCAGCACTCTGAAGACCTGACACACAGATGAGACTTCTCGAGTGGCGCTCGAGGAC 1719
2b 2800 CTCAGCACTCTGAAGACCTGACACACAGATGAGACTTCTCGAGTGGCGCTCGAGGAC 2859
2y 1720 CGAGTCAGGAGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTT 1779
2b 2860 CGAGTCAGGAGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTT 2919
2y 1780 TCACGCTCTGTCAGGCTCTCGGAGAGAGCCATCTCCGCAAAACAGTGGCTTACTAT 1839
2b 2920 TCACGCTCTGTCAGGCTCTCGGAGAGAGCCATCTCCGCAAAACAGTGGCTTACTAT 2979
2y 1840 ATCAACACAGAGACTCAACAACTTCTCGGAGAGAGCCATCTCCGCAAAACAGTGGCTTACTAT 1899
2b 2980 ATCAACACAGAGACTCAACAACTTCTCGGAGAGAGCCATCTCCGCAAAACAGTGGCTTACTAT 3039
2y 1900 TCTTTAGTGAAGCTGAATTAATGTCAGATCTCAGCTTATAGAGTGCATGAATCTCGA 1959
2b 3040 TCTTTAGTGAAGCTGAATTAATGTCAGATCTCAGCTTATAGAGTGCATGAATCTCGA 3099
2y 1960 AGACTCGAGAGCCCTTGGTGGATCTCTGAGCTCTGAGCTGCAATGATGCTGCTG 2019
2b 3100 AGACTCGAGAGCCCTTGGTGGATCTCTGAGCTCTGAGCTGCAATGATGCTGCTG 3159
2y 2020 GACGAGCACAACCTCAAGCAAAATGACCCGATGATATCTCGAGATTTAATTTGT 2079
2b 3160 GACGAGCACAACCTCAAGCAAAATGACCCGATGATATCTCGAGATTTAATTTGT 3219
2y 2080 TTGACCACTATTTATGACGCTCGAGAGAGAGCACAATTTGGTCAAGCTTCCCTCTC 2139
2b 3220 TTGACCACTATTTATGACGCTCGAGAGAGAGCACAATTTGGTCAAGCTTCCCTCTC 3279
2y 2140 TGGCTGGATATGCTGTAACCTGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGG 2199
2b 3280 TGGCTGGATATGCTGTAACCTGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGG 3339
2y 2200 ATCCGTGCTGCTTTTAAACTGCGATCATTTCCCTGTGTAAGCACAATTTGGAAGAC 2259
2b 3340 ATCCGTGCTGCTTTTAAACTGCGATCATTTCCCTGTGTAAGCACAATTTGGAAGAC 3399
2y 2260 AGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACAGGCGAGG 2319
2b 3400 AGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACAGGCGAGG 3459
2y 2320 CTGGGCTCTCTCTGATGATTTATCCAAATTCAGAGCAGTGGGTGAAGTTGATCC 2379
2b 3460 CTGGGCTCTCTCTGATGATTTATCCAAATTCAGAGCAGTGGGTGAAGTTGATCC 3519
2y 2380 TTTGGGGGAGTAACTTGGC 2401
2b 3520 TTTGGGGGAGTAACTTGGC 3541

RESULT 6

AAD37256

ID AAD37256 standard; DNA; 4966 BP.

XX

AC AAD37256;

XX DT 21-AUG-2002 (first entry)
XX DE Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
XX KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX KW Becker muscular dystrophy; ds.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Unidentified.
XX FN WO200183695-A2.
XX PD 08-NOV-2001.
XX PF 27-APR-2001; 2001WO-US13677.
XX XX 28-APR-2000; 2000US-200777P.
XX PA (XIAO/) XIAO X.
XX PI Xiao X;
XX DR WPI; 2002-049342/06.
XX XX
XX PT New dystrophin minigene for treating Duchenne or Becker muscular
XX PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX PT dystrophin gene -
XX PS Example 1; Page 59-60; 71pp; English.
XX CC The present invention relates to an isolated nucleotide sequence encoding
XX CC a dystrophin minigene. The minigene comprises N-terminal or modified
XX CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX CC invention also relates to a recombinant adeno-associated virus (AAV)
XX CC comprising dystrophin minigene operably linked to an expression control
XX CC element. The dystrophin minigene in operable linkage with an expression
XX CC control element, in a recombinant adeno-associated virus or retrovirus is
XX CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX CC vector plasmid construct containing human dystrophin minigenes, a muscle
XX CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX SQ Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;

Query Match 93.7%; Score 2250; DB 24; Length 4966;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;
QY 1 GGCAGTTCATTTGATGGAGAGTGAAGTAACTGACCTGACCGTTATCAACAGCTTTAGAGAA 60
Db 1757 GGCAGTTCATTTGATGGAGAGTGAAGTAACTGACCGTTATCAACAGCTTTAGAGAA 1816
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGACACCAAGAGAGATTTCTAAT 120
Db 1817 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGACACCAAGAGAGATTTCTAAT 1876
QY 121 GATGTGAAGTGGTGAAGACCACTGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
Db 1877 GATGTGAAGTGGTGAAGACCACTGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 1936
QY 181 GCCCATCAGCGCCGGTGGTGAATATTTCTACAAATTCGGAGTAAAGCTGATGGAACAGGA 240
Db 1937 GCCCATCAGCGCCGGTGGTGAATATTTCTACAAATTCGGAGTAAAGCTGATGGAACAGGA 1996
QY 241 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 300
Db 1997 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 2056
QY 301 TGGGAATGCCTCAGGCTAGCTAGCTAGTGAAGAAACAAAGCAATTTACATAGATTTAATG 360

Db 2057 TGGGAATGCTCAGGGTAGCTAGCTAGGAAAAAACAAGCAATTTACATAGAGTTTAAATG 2116
Qy 361 GATCTCCAGAAATCAGAACTGAAAGTTGAATGATGCTGGCTTAACAAAACAGAGAAAGA 420
Db 2117 GATCTCCAGAAATCAGAACTGAAAGTTGAATGATGCTGGCTTAACAAAACAGAGAAAGA 2176
Qy 421 ACAAGAAAATGGAGGAAGAGCTCTTTGAGCCTGATCTTTGAAGACTTAAACGCCCAAGTA 480
Db 2177 ACAAGAAAATGGAGGAAGAGCTCTTTGAGCCTGATCTTTGAAGACTTAAACGCCCAAGTA 2236
Qy 481 CAAACAATAGGTGCTTCAAGAAGATCTAGAAACAGAAACAAAGTCAGGGTCAATTTCTTC 540
Db 2237 CAAACAATAGGTGCTTCAAGAAGATCTAGAAACAGAAACAAAGTCAGGGTCAATTTCTTC 2296
Qy 541 ACTCAGATGGTGGTGTAGTTGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGAA 600
Db 2297 ACTCAGATGGTGGTGTAGTTGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGAA 2356
Qy 601 GAACAATTAAGGTATGGAGATCGATGGCAACATCTGATGAGACAGAAACCGC 660
Db 2357 GAACAATTAAGGTATGGAGATCGATGGCAACATCTGATGAGACAGAAACCGC 2416
Qy 661 TGGTCTTTTACAGAC----- 678
Db 2417 TGGTCTTTTACAGACCGAGCTGACCTAGCTCCTGGAGTGAACCACTATTGGAGCCTCT 2476
Qy 679 ----- 678
Db 2477 CTTACTCAGACTTACTCTGGTGACAAACCTGGTTACTTAAGGAAACTGCCATCTCC 2536
Qy 679 ----- 678
Db 2537 AAACCTAGAAATGCCATCTTCCTGTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 2596
Qy 700 TTCCCTCCGACCTGAAAAGTTTCTTGCTGGCTTACAGAGCTGAAACAACTGCCAAT 759
Db 2597 TTCCCTCCGACCTGAAAAGTTTCTTGCTGGCTTACAGAGCTGAAACAACTGCCAAT 2656
Qy 760 GTCCCTACAGGATGCTACCGTAAAGAAAGCTCTAGAGACTCCAGGGAGTAAAGAG 819
Db 2657 GTCCCTACAGGATGCTACCGTAAAGAAAGCTCTAGAGACTCCAGGGAGTAAAGAG 2716
Qy 820 CTGATGAACAATGGAAGACCTCCAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 879
Db 2717 CTGATGAACAATGGAAGACCTCCAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 2776
Qy 880 AACCTGGATGAACAGCCAAATAATCTAGATCCCTGGAGGTCCGATGATGCAATC 939
Db 2777 AACCTGGATGAACAGCCAAATAATCTAGATCCCTGGAGGTCCGATGATGCAATC 2836
Qy 940 CTGTTACAAAGACGTTTGGATTAACATGACTCAAGTGGAGTGAATTCGGAAAGTCT 999
Db 2837 CTGTTACAAAGACGTTTGGATTAACATGACTCAAGTGGAGTGAATTCGGAAAGTCT 2896
Qy 1000 CTCACATTTAGTCCCATTTGGAAGCCAGTCTGACCAAGTGAAGCTGCACTCTCT 1059
Db 2897 CTCACATTTAGTCCCATTTGGAAGCCAGTCTGACCAAGTGAAGCTGCACTCTCT 2956
Qy 1060 CTGAGGAATCTTGGTGGCTACAGCTGAAGAATGATGAATTAAGCCGGCAGGACCT 1119
Db 2957 CTGAGGAATCTTGGTGGCTACAGCTGAAGAATGATGAATTAAGCCGGCAGGACCT 3016
Qy 1120 ATTGGAGGCGACTTTCACAGTTCAGAAAGCAGAGATGATACATAGGCTTCAAGAGG 1179
Db 3017 ATTGGAGGCGACTTTCACAGTTCAGAAAGCAGAGATGATACATAGGCTTCAAGAGG 3076
Qy 1180 GAATTAAGAACTAAAGAACTGTAATCATAGATCTCTTGAAGTGTACGAATTTCTG 1239
Db 3077 GAATTAAGAACTAAAGAACTGTAATCATAGATCTCTTGAAGTGTACGAATTTCTG 3136
Qy 1240 ACAGAGCCCTTTGGAAGGCTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCT 1299

Db 3137 ACAGAGCAGCCTTTGCAAGGACTAGAGAAACTCTACCAGAGCCACAGAGAGCTGCTCT 3196
Qy 1300 GAGGAGAGAGCCAGAAATGTCACTCGGCTTTACGAAACAGGCTGAGGAGTCAATACT 1359
Db 3197 GAGGAGAGAGCCAGAAATGTCACTCGGCTTTACGAAACAGGCTGAGGAGTCAATACT 3256
Qy 1360 GAGTGGGAAAAATTTGAACCTTGCACTCGGCTGACTGGCAGAGAAAAATAGATGAGACCTTT 1419
Db 3257 GAGTGGGAAAAATTTGAACCTTGCACTCGGCTGACTGGCAGAGAAAAATAGATGAGACCTTT 3316
Qy 1420 GAAAGACTCCAGAACTTCAAGAGGCCACGATGAGCTGACCTCAAGCTGCGCAAGCT 1479
Db 3317 GAAAGACTCCAGAACTTCAAGAGGCCACGATGAGCTGACCTCAAGCTGCGCAAGCT 3376
Qy 1480 GAGGTGATCAAGGATCTCTGGCAGCCCTGGGCGATCTCTCAATGACTCTCTCCAAAGAT 1539
Db 3377 GAGGTGATCAAGGATCTCTGGCAGCCCTGGGCGATCTCTCAATGACTCTCTCCAAAGAT 3436
Qy 1540 CACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGCGCTCTGAAAGAGACGTGAGC 1599
Db 3437 CACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGCGCTCTGAAAGAGACGTGAGC 3496
Qy 1600 CAGCTCAATGACCTTCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 1659
Db 3497 CAGCTCAATGACCTTCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 3556
Qy 1660 CTGAGCACTCTGAAAGACCTTGAACACAGATGGAAGCTTCTGAGGTGGCCCTGAGGAC 1719
Db 3557 CTGAGCACTCTGAAAGACCTTGAACACAGATGGAAGCTTCTGAGGTGGCCCTGAGGAC 3616
Qy 1720 CGAGTCAGGAGCTGATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 1779
Db 3617 CGAGTCAGGAGCTGATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTT 3676
Qy 1780 TCCACGCTCTGTCCAGGTCCTCTGGGAGAGCCATCTGCCAAAACAAAGTGGCCCTACTAT 1839
Db 3677 TCCACGCTCTGTCCAGGTCCTCTGGGAGAGCCATCTGCCAAAACAAAGTGGCCCTACTAT 3736
Qy 1840 ATCAACACAGAGACTCAAAACAACTTGTGGGAGACCATCTGCCAAAACAAAGTGGCCCTACTAC 1899
Db 3737 ATCAACACAGAGACTCAAAACAACTTGTGGGAGACCATCTGCCAAAACAAAGTGGCCCTACTAC 3796
Qy 1900 TCTTTAGCTGACCTGAATTAATGTGAGATCTCAGCTTATAGGACTGCCATGAACTCCGA 1959
Db 3797 TCTTTAGCTGACCTGAATTAATGTGAGATCTCAGCTTATAGGACTGCCATGAACTCCGA 3856
Qy 1960 AGACTCAGAGAGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGATGATGCTCTG 2019
Db 3857 AGACTCAGAGAGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGATGATGCTCTG 3916
Qy 2020 GACCAGCAACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATTATTAATTGT 2079
Db 3917 GACCAGCAACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATTATTAATTGT 3976
Qy 2080 TTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTC 2139
Db 3977 TTGACCACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTC 4036
Qy 2140 TGGTGGATATGCTGCTGAACCTGGCTGCTGAATTTATGATACGGAGCAACAGGAGG 2199
Db 4037 TGGTGGATATGCTGCTGAACCTGGCTGCTGAATTTATGATACGGAGCAACAGGAGG 4096
Qy 2200 ATCCGTGCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGAC 2259
Db 4097 ATCCGTGCTCTGCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGAC 4156
Qy 2260 AAGTACAGATACCTTTTCAAGAGTGGCAAGTTCAGCAGGATTTTGTACAGCGCAGG 2319
Db 4157 AAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAGCAGGATTTTGTACAGCGCAGG 4216
Qy 2320 CTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCAAAGACAGTGGGTGAAGTTGATCC 2379
Db 4217 CTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCAAAGACAGTGGGTGAAGTTGATCC 4276

QY 2380 TTGCGGGGAGTAACATTGAGC 2401
DB 4277 TTGCGGGGAGTAACATTGAGC 4298

RESULT 7
AAD37262
ID AAD37262 standard; DNA; 4990 BP.

XX AAD37262;

DT 21-AUG-2002 (first entry)

XX Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.

XX Chimeric - Homo sapiens.

DS Chimeric - Cytomegalovirus.

DS Chimeric - Unidentified.

XX WO200183695-A2.

PN 08-NOV-2001.

PD 27-APR-2001; 2001WO-US13677.

PR 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/05.

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

XX Example 1; Page 67-68; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomagalovirus (CMV) promoter and a small polyA signal sequence.

XX Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 other;

Query Match 93.7%; Score 2250; DB 24; Length 4990;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGTTCATTGATGGAGTGAAGTAACCTGGACCGTATCAACACGCTTTAGAGAA 60
DB 1781 GGCAGTTCATTGATGGAGTGAAGTAACCTGGACCGTATCAACACGCTTTAGAGAA 1840

QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 120

DB 1841 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 1900

QY 121 GATGTGAAGTGGTGAAGAACACCGATTTCATCTATGAGGGGTACATGATGATTGACA 180

DB 1901 GATGTGAAGTGGTGAAGAACACCGATTTCATCTATGAGGGGTACATGATGATTGACA 1960
QY 181 GCCCATCAGGGCCGGGTTGGTAATATCTACAAATTTGGAGTAAGCTGATTGGAACAGGA 240
DB 1961 GCCCATCAGGGCCGGGTTGGTAATATCTACAAATTTGGAGTAAGCTGATTGGAACAGGA 2020
QY 241 AAATTATCAGAAGATGAAGAAACTGGAAGTACAAGAGCAGATGAATCTCTTAAATTCAGA 300
DB 2021 AAATTATCAGAAGATGAAGAAACTGGAAGTACAAGAGCAGATGAATCTCTTAAATTCAGA 2080
QY 301 TGGGAATGCTCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
DB 2081 TGGGAATGCTCTCAGGCTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 2140
QY 361 GATCTCCAGAAATCAGAAACTGAAAAGAGTTGAAAGACTGGCTTAAACAAAAACAGAAAGA 420
DB 2141 GATCTCCAGAAATCAGAAACTGAAAAGAGTTGAAAGACTGGCTTAAACAAAAACAGAAAGA 2200
QY 421 ACAAGAAAAATGGAGGAAGAGCCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTA 480
DB 2201 ACAAGAAAAATGGAGGAAGAGCCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAAGTA 2260
QY 481 CACCAACATAGGTGCTTCAAGAGATCTAGAACAGACCAAGTCAAGGTCAATTTCTCTC 540
DB 2261 CACCAACATAGGTGCTTCAAGAGATCTAGAACAGACCAAGTCAAGGTCAATTTCTCTC 2320
QY 541 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
DB 2321 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2380
QY 601 GAACAACCTTAAGTATTTGGGAGATCGATGGGCAACATCTGTAGTGGACAGAGACCCG 660
DB 2381 GAACAACCTTAAGTATTTGGGAGATCGATGGGCAACATCTGTAGTGGACAGAGACCCG 2440
QY 661 TGGGTTCCTTTTACAAGAC----- 678
DB 2441 TGGGTTCCTTTTACAAGACCGCTGACCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 2500
QY 679 ----- 678
DB 2501 CCTACTCAGACTGTTACTCTGTGTGACACACCTGTGTTACTAAGGAAACTGCCATCTCC 2560
QY 679 -----ACTCATAGATTACTGCAACAG 699
DB 2561 AAACAGAAATGCCATCTTCTGTGATGTGGAGGTACCTACTACTAGATTACTGCAACAG 2620
QY 700 TTCCCTCTGGACCTGGAAAAAGTTTCTTGGCTTACAGAACTGAAACAACTGCCAAT 759
DB 2621 TTCCCTCTGGACCTGGAAAAAGTTTCTTGGCTTACAGAACTGAAACAACTGCCAAT 2680
QY 760 GTCTACAGATGCTACCCCTAGGAAGGCTCTAGAGACTCCAGGAGTAAAGAG 819
DB 2681 GTCTACAGATGCTACCCCTAGGAAGGCTCTAGAGACTCCAGGAGTAAAGAG 2740
QY 820 CTGATGAAACAAATGGCAAGACCTCCCAAGGTGAATGAAGTCAACACAGATGTTTATCAC 879
DB 2741 CTGATGAAACAAATGGCAAGACCTCCCAAGGTGAATGAAGTCAACACAGATGTTTATCAC 2800
QY 880 AACCTGGATGAAACAGCCAAAATCTCGAGATCCCTGGAAGTCCGATGATGAGTC 939
DB 2801 AACCTGGATGAAACAGCCAAAATCTCGAGATCCCTGGAAGTCCGATGATGAGTC 2860
QY 940 CTGTTACAAAGAGCTTTGGGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCT 999
DB 2861 CTGTTACAAAGAGCTTTGGGATAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCT 2920
QY 1000 CTCAACATTAGTCCCATTTTGGAGGACAGTTCTGACAGTGGAGCGGTCTGCAACCTTTCT 1059
DB 2921 CTCAACATTAGTCCCATTTTGGAGGACAGTTCTGACAGTGGAGCGGTCTGCAACCTTTCT 2980
QY 1060 CTGCAGGAATCTTGTGTGCTTACAGCTGAAGATGATGAATTAAGCGCGGAGGACCT 1119

2981 CTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGACACCT 3040
1120 ATTGAGGCGATTTCCAGCAGTTCCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGG 1179
3041 ATTGAGGCGATTTCCAGCAGTTCCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGG 3100
1180 GAATTTGAAACCTTAAAGAACCTTGTATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG 1239
3101 GAATTTGAAACCTTAAAGAACCTTGTATCATGAGTACTCTTTGAGACTGTACGAATATTTCTG 3160
1240 ACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGGAGCCGAGAGCTGCTCCT 1299
3161 ACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGGAGCCGAGAGCTGCTCCT 3220
1300 GAGGAGAGAGCCAGAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACT 1359
3221 GAGGAGAGAGCCAGAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACT 3280
1360 GAGTGGGAAAATTTGAACCTGCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCTTT 1419
3281 GAGTGGGAAAATTTGAACCTGCACTCCGCTGACTGGCGAGAGAAAATAGATGAGACCTTT 3340
1420 GAAAGACTCCAGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTCGGCGAAGCT 1479
3341 GAAAGACTCCAGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTCGGCGAAGCT 3400
1480 GAGGTGATCAAGGATCTTGGCAGCCGCTGGCGGATCTCTCAATGACTCTCTCAAGAT 1539
3401 GAGGTGATCAAGGATCTTGGCAGCCGCTGGCGGATCTCTCAATGACTCTCTCAAGAT 3460
1540 CACCTCGAAGAGTCAAGCAGCTTCAAGAGGAAATGGCCCTCGAAGAGAAAGTGAGC 1599
3461 CACCTCGAAGAGTCAAGCAGCTTCAAGAGGAAATGGCCCTCGAAGAGAAAGTGAGC 3520
1600 CAGCTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCACTTCAAGCTCTCAAGTATAAC 1659
3521 CAGCTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCACTTCAAGCTCTCAAGTATAAC 3580
1660 CTCAGACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGTGGAGAC 1719
3581 CTCAGACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGTGGAGAC 3640
1720 CGAGTCAGGACGCTGATCAAGCCACAGAGGACTTGGTCCAGCATCTCAGACCTTTCTT 1779
3641 CGAGTCAGGACGCTGATCAAGCCACAGAGGACTTGGTCCAGCATCTCAGACCTTTCTT 3700
1780 TCCAGCTCTGTCAGGCTCCCTGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTTACTAT 1839
3701 TCCAGCTCTGTCAGGCTCCCTGGAGAGAGCCATCTCGCCAAAACAAAGTGCCCTTACTAT 3760
1840 ATCAACCCAGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAG 1899
3761 ATCAACCCAGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAG 3820
1900 TCTTTAGCTGACCTGAATATGATGATCTCAGCTTATAGGACTGCCATGAATCCGA 1959
3821 TCTTTAGCTGACCTGAATATGATGATCTCAGCTTATAGGACTGCCATGAATCCGA 3880
1960 AGACTCAGAGGCGCTTGTGGTGTCTTGGAGCTTCTGAGCTTCTGAGCTGATGATGCTT 2019
3881 AGACTCAGAGGCGCTTGTGGTGTCTTGGAGCTTCTGAGCTTCTGAGCTGATGATGCTT 3940
2020 GACCAGCAACACCTCAAGCAAAATGACAGCCGCTGATATCTCAGATTTATTAATGT 2079
3941 GACCAGCAACACCTCAAGCAAAATGACAGCCGCTGATATCTCAGATTTATTAATGT 4000
2080 TTGACACATTTATGACGCTGAGAGAGAGACACAAATTTGGTCAAGCTCCTCTC 2139
4001 TTGACACATTTATGACGCTGAGAGAGAGACACAAATTTGGTCAAGCTCCTCTC 4060
2140 TGGTGGATATGTCTGAACTGGCTGCTGAATTTATGATACGGAGCAAGAGGAGG 2199
4061 TGGTGGATATGTCTGAACTGGCTGCTGAATTTATGATACGGAGCAAGAGGAGG 4120

2200 ATCCGCTCTCTCTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 2259
4121 ATCCGCTCTCTCTCTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGAC 4180
2260 AGATACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGG 2319
4181 AGATACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGG 4240
2320 CTGGGCTCTCTCTCTGATCATTTCTATCCAAATTTCAAGACAGTTCGGTGAAGTTGCATCC 2379
4241 CTGGGCTCTCTCTCTGATCATTTCTATCCAAATTTCAAGACAGTTCGGTGAAGTTGCATCC 4300
2380 TTGCGGGCAGTAACATTGAGC 2401
4301 TTGCGGGCAGTAACATTGAGC 4322

RESULT 8

AAD37230
ID AAD37230 standard; DNA; 4182 BP.
XX
AC AAD37230;
XX
DT 21-AUG-2002 (first entry)
XX Human dystrophin minigene delta4173.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
FN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI XIAO X;
XX
DR WFI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
FS Example 1; Page 43-44; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is human
XX dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
XX hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
XX hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 other;

Query Match 86.1%; Score 2067; DB 24; Length 4182;
Best Local Similarity 88.1%; Pred. No. 0;

Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

```
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTCGACCGTTTATCAAAACAGCTTTAGAGAA 60
DB 1000 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTCGACCGTTTATCAAAACAGCTTTAGAGAA 1059
2Y 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT 120
DB 1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTTCTAAT 1119
QY 121 GATGTGGAAGTGTGAAAGACCAAGTTTCTACTCATGAGGGGTACATGATGATTTGACA 180
DB 1120 GATGTGGAAGTGTGAAAGACCAAGTTTCTACTCATGAGGGGTACATGATGATTTGACA 1179
QY 181 GCCCATCAGGGCCGGTGGTAAATTTCTACAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 240
DB 1180 GCCCATCAGGGCCGGTGGTAAATTTCTACAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 1239
2Y 241 AAATTATCAGAAGATGAGAAACTGAAAGTACAGACAGATGAATCTCTAAATTCAGA 300
DB 1240 AAATTATCAGAAGATGAGAAACTGAAAGTACAGACAGATGAATCTCTCTAAATTCAGA 1299
2Y 301 TGGGAATGCTCAGGGTAGCTAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
DB 1300 TGGGAATGCTCAGGGTAGCTAGCATGGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1359
2Y 361 GATCTCCAGAACTCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAACAGAGAAAGA 420
DB 1360 GATCTCCAGAACTCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAACAGAGAAAGA 1419
2Y 421 ACAAGGAAATGAGGAAGAGCTCTTGACCTGATCTTTGAGAGCTTAAAGCCCAAGTA 480
DB 1420 ACAAGGAAATGAGGAAGAGCTCTTGACCTGATCTTTGAGAGCTTAAAGCCCAAGTA 1479
2Y 481 CAACAACAATAAGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTCTC 540
DB 1480 CAACAACAATAAGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTCTC 1539
2Y 541 ACTCACATGGTGGTGTAGTGTGATGATCTAGTGGAGATCAGCAACTGCTGCTTTGAA 600
DB 1540 ACTCACATGGTGGTGTAGTGTGATGATCTAGTGGAGATCAGCAACTGCTGCTTTGAA 1599
2Y 601 GAACAACCTTAAAGTATGGGAGATCGATGGGAAACATCTGTAGATGACAGAACACCGC 660
DB 1600 GAACAACCTTAAAGTATGGGAGATCGATGGGAAACATCTGTAGATGACAGAACACCGC 1659
2Y 661 TGGGTTCTTTTCAAGAC----- 678
DB 1660 TGGGTTCTTTTCAAGACATCTTCTCAAAATGGCAACGTTTACTGAAAGAACAGTGCCTT 1719
2Y 679 ----- 678
DB 1720 TTTAGTGATGGCTTTCAGAAAAAGAGATGCAGTGAACAAGATTCACAACTGGCTTTT 1779
2Y 679 ----- 678
DB 1780 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAATGCGCGTTTTTAAAGCGGATCTA 1839
2Y 679 ----- 678
DB 1840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACNAGATCTTTTTCACAA 1899
2Y 679 ----- 678
DB 1900 CTGAAGAATAAGTCAGTGACCCAGAGACGGAGCATGCTGGTAACTTTGCGCGGTGT 1959
2Y 679 ----- 678
DB 1960 TGGGATAAATTTAGTCCAAAAAATTGAAAAAGATACAGCACAGACTCATAGATTTACTGCAA 2019
2Y 697 CAGTTCCTCCCTGACCTGAAAGATTTCTTGCTGCTTACAGAGAGCTGAAACAACTGCC 756
DB 2020 CAGTTCCTCCCTGACCTGAAAGATTTCTTGCTGCTTACAGAGAGCTGAAACAACTGCC 2079
```

```
QY 757 AATGCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGAGTAAAA 816
DB 2080 AATGCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAAGGAGTAAAA 2139
QY 817 GAGCTGATGAAACAATGGCAAGACCTCAAAGGTGAAATTTGAAAGTCAACAGATGTTTAT 876
DB 2140 GAGCTGATGAAACAATGGCAAGACCTCAAAGGTGAAATTTGAAAGTCAACAGATGTTTAT 2199
QY 877 CACAACCTGGATGAAACACGCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 936
DB 2200 CACAACCTGGATGAAACACGCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCA 2259
QY 937 GTCTGTTTACAAAGACGTTTGGATAAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAG 996
DB 2260 GTCTGTTTACAAAGACGTTTGGATAAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAG 2319
QY 997 TCTCTCAACATTAGTCCCATTTTGGAGCCAGTCTGACCAGTGGAGGCTCTGCACCTT 1056
DB 2320 TCTCTCAACATTAGTCCCATTTTGGAGCCAGTCTGACCAGTGGAGGCTCTGCACCTT 2379
QY 1057 TCTCTCAGAGAACTTCTGCTGTGCTACAGCTGAAAAGATGATGAATTAAGCCGCGCAGGCA 1116
DB 2380 TCTCTCAGAGAACTTCTGCTGTGCTACAGCTGAAAAGATGATGAATTAAGCCGCGCAGGCA 2439
QY 1117 CTTATTGGAGGCGACTTTCAGCAGTTCAGAGCAGAACGATGTACATAGGGCTTTCAAG 2499
DB 2440 CTTATTGGAGGCGACTTTCAGCAGTTCAGAGCAGAACGATGTACATAGGGCTTTCAAG 2499
QY 1177 AGGAATTTGAAAACTTAAAGAACCTGTAAATCATCAGTACTCTTCAGACTGTACGAATATTT 1236
DB 2500 AGGAATTTGAAAACTTAAAGAACCTGTAAATCATCAGTACTCTTCAGACTGTACGAATATTT 2559
QY 1237 CTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGTGCCT 1296
DB 2560 CTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGTGCCT 2619
QY 1297 CCTGAGGAGAGCCAGATGTCACCTCGCTTCTAGGAAGCAGGCTGAGGAGTCAAT 1356
DB 2620 CCTGAGGAGAGCCAGATGTCACCTCGCTTCTAGGAAGCAGGCTGAGGAGTCAAT 2679
QY 1357 ACTGAGTGGGAAAAATTTGAACCTCGCTACTCGCTGAGTGGCAGAGAAAAATAGATGAGACC 1416
DB 2680 ACTGAGTGGGAAAAATTTGAACCTCGCTACTCGCTGAGTGGCAGAGAAAAATAGATGAGACC 2739
QY 1417 CTTGAAAGATCTCAGGAACTTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCAA 1476
DB 2740 CTTGAAAGATCTCAGGAACTTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCAA 2799
QY 1477 GCTGAGTGTATCAAGGGATCCTGCGCAGCCGCTGGCGATCTCTCATTCACATCTCTCCAA 1536
DB 2800 GCTGAGTGTATCAAGGGATCCTGCGCAGCCGCTGGCGATCTCTCATTCACATCTCTCCAA 2859
QY 1537 GATCAGCTCGAGAAAGTCAAGGCACTTTCAGAGGAGAAATTTGCGCTCTGAAAGAGAACGTTG 1596
DB 2860 GATCAGCTCGAGAAAGTCAAGGCACTTTCAGAGGAGAAATTTGCGCTCTGAAAGAGAACGTTG 2919
QY 1597 AGCCAGTCAATGACCTTCTGCGCAGCTTACACCTTGGGATTCAGCTCTCACCGTAT 1656
DB 2920 AGCCAGTCAATGACCTTCTGCGCAGCTTACACCTTGGGATTCAGCTCTCACCGTAT 2979
QY 1657 AACTCAGCACTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTTGGCGCTCGAG 1716
DB 2980 AACTCAGCACTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTTGGCGCTCGAG 3039
QY 1717 GACCGAGTCAAGGAGCTGATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTT 1776
DB 3040 GACCGAGTCAAGGAGCTGATGAAGCCACAGGGAATTTGGTCCAGCATCTCAGCACTTT 3099
QY 1777 CTTTCCACAGTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTAC 1836
DB 3100 CTTTCCACAGTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTAC 3159
```

QY 1837 TATATCAACCCAGGAGCTCAAACTTCTGGGACCATCCAAAATGACAGAGCTTAC 1896
DB |||||
QY 3160 TATATCAACCCAGGAGCTCAAACTTCTGGGACCATCCAAAATGACAGAGCTTAC 3219
DB |||||
QY 1897 CAGTCTTAGCTGAGCTGAATAATCTCAGTCTTAGCTGAGCTGAGCTGAGCTG 1956
DB |||||
QY 3220 CAGTCTTAGCTGAGCTGAATAATCTCAGTCTTAGCTGAGCTGAGCTGAGCTG 3279
DB |||||
QY 1957 CCAAGAGCTGACAGAGCCCTTCTGCTGATCTCTTGAGCTGCTGAGCTGAGCTG 2016
DB |||||
QY 3280 CCAAGAGCTGACAGAGCCCTTCTGCTGATCTCTTGAGCTGCTGAGCTGAGCTG 3339
DB |||||
QY 2017 TTGGACAGCAACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAAT 2076
DB |||||
QY 3340 TTGGACAGCAACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAAT 3399
DB |||||
QY 2077 TGTCTTGACACATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAGTCCCT 2136
DB |||||
QY 3400 TGTCTTGACACATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAGTCCCT 3459
DB |||||
QY 2137 CTCTGGTGGATATGTCTGAACTGGCTGCTGAAATGTTATGATACGGGACGAGGG 2196
DB |||||
QY 3460 CTCTGGTGGATATGTCTGAACTGGCTGCTGAAATGTTATGATACGGGACGAGGG 3519
DB |||||
QY 2197 AGGATCCGTGCTCTCTTTTAAACTGGCATCTTTCCCTGTGTAAAGCAATTTGAA 2256
DB |||||
QY 3520 AGGATCCGTGCTCTCTTTTAAACTGGCATCTTTCCCTGTGTAAAGCAATTTGAA 3579
DB |||||
QY 2257 GACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACAGGC 2316
DB |||||
QY 3580 GACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACAGGC 3639
DB |||||
QY 2317 AGGCTGGGCTCTCTCTGATGATCTATTCMAATTCACAGAGTGGGTGAAAGTTGCA 2376
DB |||||
QY 3640 AGGCTGGGCTCTCTCTGATGATCTATTCMAATTCACAGAGTGGGTGAAAGTTGCA 3699
DB |||||
QY 2377 TCTTTTGGGGGAGTAACATTGAGC 2401
DB |||||
QY 3700 TCTTTTGGGGGAGTAACATTGAGC 3724
DB |||||

RESULT 9
AAD37255
ID AAD37255 standard; DNA; 5149 BP.
XX AC AAD37255;
XX DT 21-AUG-2002 (first entry)
XX DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX KW Becker muscular dystrophy; ds.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Unidentified.
XX PN WO200183695-A2.
XX PD 08-NOV-2001.
XX PF 27-APR-2001; 2001WO-US13677.
XX PR 28-APR-2000; 2000US-200777P.
XX PA (XIAO/) XIAO X.
XX PI Xiao X;
XX DR WPI; 2002-049342/06.
XX PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

XX Example 1; Page 57-59; 7lpp; English.

PS The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.

XX Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 other;

QY Query Match 86.1%; Score 2067; DB 24; Length 5149;
DB Best Local Similarity 88.1%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

QY 1 GGCAAGTTCATTGATGAGAGTGAAGTAACTGACCTGGACCTTATCAACAGCTTTAGAAGA 60
DB 1757 GGCAAGTTCATTGATGAGAGTGAAGTAACTGACCTGGACCTTATCAACAGCTTTAGAAGA 1816
QY 61 GTATTATCTGTGGCTTCTTCTCTGAGGACACATTTGCAAGCAACAAGGAGAGATTTCTAAT 120
DB 1817 GTATTATCTGTGGCTTCTTCTCTGAGGACACATTTGCAAGCAACAAGGAGAGATTTCTAAT 1876
QY 121 GATGTGGAAGTGGTGAAGACAGATTTCTATCTATGAGGGGTACATGATGATTTGACA 180
DB 1877 GATGTGGAAGTGGTGAAGACAGATTTCTATCTATGAGGGGTACATGATGATTTGACA 1936
QY 181 GCCCATCAGGCGCGGTTCGTATATTTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGA 240
DB 1937 GCCCATCAGGCGCGGTTCGTATATTTCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGA 1996
QY 241 AAATTATCAGAAAGTGAAGAACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGA 300
DB 1997 AAATTATCAGAAAGTGAAGAACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGA 2056
QY 301 TGGGAATGCTCAGGGTACGTAGCATGGAAGAAACCAAGCAATTTACATAGAGTCTTAAATG 360
DB 2057 TGGGAATGCTCAGGGTACGTAGCATGGAAGAAACCAAGCAATTTACATAGAGTCTTAAATG 2116
QY 361 GATCTCCAGAACTCAGAAAGTGAAGTGAATGATCTGGCTAACAAAAACAGAAAGA 420
DB 2117 GATCTCCAGAACTCAGAAAGTGAAGTGAATGATCTGGCTAACAAAAACAGAAAGA 2176
QY 421 ACAAGGAAATGAGGAAGAGCTCTTGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
DB 2177 ACAAGGAAATGAGGAAGAGCTCTTGACCTGATCTTGAAGACCTTAAACGCCAAGTA 2236
QY 481 CAACAAACATAAGGTGCTTCAAGAAAGATCTAGAACAGAAACAAGTCAAGGTCTCAATCTCTC 540
DB 2237 CAACAAACATAAGGTGCTTCAAGAAAGATCTAGAACAGAAACAAGTCAAGGTCTCAATCTCTC 2296
QY 541 ACTCAATGCTGGTGTAGTGTAGTGAATCTAGTGGAGATCAGCAACTGCTCTCTTTGGA 600
DB 2297 ACTCAATGCTGGTGTAGTGTAGTGAATCTAGTGGAGATCAGCAACTGCTCTCTTTGGA 2356
QY 601 GAACAACTTAAGGTATTGGGAGATCCATGGCAACATCTGTAGATGGACAGAACCCGC 660
DB 2357 GAACAACTTAAGGTATTGGGAGATCCATGGGCAACATCTGTAGATGGACAGAACCCGC 2416
QY 661 TGGGTTCTTTTACAAGAC----- 678
DB 2417 TGGGTTCTTTTACAAGACCTTCTCAAAATGCAACAGCTTACTGAGAGACAGTGCCTT 2476

2Y 679 ----- 678
Db 2477 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGTCAGTGAACAAGATTCACACAACTGGCTTT 2536
2Y 679 ----- 678
Db 2537 AAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGGCGTTTAAAGCGGATCTA 2596
2Y 679 ----- 678
Db 2597 GAAAGAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACA 2656
2Y 679 ----- 678
Db 2657 CTGAAGATAAGTCAAGTGACCCAGAAAGACGGAAGCATGGCTGGATAAATTTTCCCGGTGT 2716
2Y 679 ----- 696
2Y 2717 TGGGATAATTAGTCCAAAAAATTGAAAAAGATACAGCACAGCTCATAGATTACTGCNA 2776
2Y 697 CAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTTACAGAAGCTGAACAACTGCC 756
2Y 2777 CAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTTACAGAAGCTGAACAACTGCC 2836
2Y 757 AATGCTCTACAGATGCTACCTGTAAGGAAGGCTCTCTAGAGACTCCAGAGGAGTAAA 816
2Y 2837 AATGCTCTACAGATGCTACCTGTAAGGAAGGCTCTCTAGAGACTCCAGAGGAGTAAA 2896
2Y 817 GAGCTGATGAAACAAATGGCAAGACCTCCAAAGTGAAATTTGAAGCTCACACAGATGTTAT 876
2Y 2897 GAGCTGATGAAACAAATGGCAAGACCTCCAAAGTGAAATTTGAAGCTCACACAGATGTTAT 2956
2Y 877 CACAACTGATGAAACAAAGCCAAAAAATCTTGAGATCCCTGGAAGTTCCGATGATGCA 936
2Y 2957 CACAACTGATGAAACAAAGCCAAAAAATCTTGAGATCCCTGGAAGTTCCGATGATGCA 3016
2Y 937 GTCCTGTTACAAAGAGCTTTGGATAACATGAACCTTCAAGTGAGTGAACTTCGGAANAAG 996
2Y 3017 GTCCTGTTACAAAGAGCTTTGGATAACATGAACCTTCAAGTGAGTGAACTTCGGAANAAG 3076
2Y 997 TCTCTCAACATTAAGTGCCCATTTGGAAAGCAGTTCTGACAGTGGAAGCTGTGACCTTT 1056
2Y 3077 TCTCTCAACATTAAGTGCCCATTTGGAAAGCAGTTCTGACAGTGGAAGCTGTGACCTTT 3136
2Y 1057 TCTCTCAGAACTTCTGCTGGCTTACAGCTGAAGATGATGAATTAAGCCGCGCAGGCA 1116
2Y 3137 TCTCTCAGAACTTCTGCTGGCTTACAGCTGAAGATGATGAATTAAGCCGCGCAGGCA 3196
2Y 1117 CCTATTGGAGCGACTTTCCAGCAGTTTCAGAAAGCAGATGTATACATAGGCGCTTCAAG 1176
2Y 3197 CCTATTGGAGCGACTTTCCAGCAGTTTCAGAAAGCAGATGTATACATAGGCGCTTCAAG 3256
2Y 1177 AGGAAATTGAAACTAAAGAACTGTAACTGATGATGATCTCTTGAGACTGTACGAATATT 1236
2Y 3257 AGGAAATTGAAACTAAAGAACTGTAACTGATGATGATCTCTTGAGACTGTACGAATATT 3316
2Y 1237 CTGACAGCAGCCTTTGGAAAGCTAGAGAACTCTACAGAGGCCAGAGAGCTGCT 1296
2Y 3317 CTGACAGCAGCCTTTGGAAAGCTAGAGAACTCTACAGAGGCCAGAGAGCTGCT 3376
2Y 1297 CTTGAGGAGAGGCCCAAGATGTCACTGGCTTTCTACGAAAGCAGGCTGAGAGGTCAAT 1356
2Y 3377 CTTGAGGAGAGGCCCAAGATGTCACTGGCTTTCTACGAAAGCAGGCTGAGAGGTCAAT 3436
2Y 1357 ACTGAGTGGGAAAAATTTGAACTTGCCTCGCTGACTGGCAGAGAAATAGATGAGCC 1416
2Y 3437 ACTGAGTGGGAAAAATTTGAACTTGCCTCGCTGACTGGCAGAGAAATAGATGAGCC 3496
2Y 1417 CTTGAAAGACTCCAGGAACCTTAAGAGGCCACAGGATGAGCTGAGCTCAAGCTGCGCCAA 1476
2Y 3497 CTTGAAAGACTCCAGGAACCTTAAGAGGCCACAGGATGAGCTGAGCTCAAGCTGCGCCAA 3556
2Y 1477 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGCTGGGAGATCTCTCATTTGACTCTCCAA 1536

Db 3557 GCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGATCTCTCTCATTTGACTCTCTCCAA 3616
QY 1537 GATCACCTCGAGAAAGTCAAGGCACCTTCAGAGGAGAAATTCGGCTCTGAAAGAGAAAGTG 1596
Db 3617 GATCACCTCGAGAAAGTCAAGGCACCTTCAGAGGAGAAATTCGGCTCTGAAAGAGAAAGTG 3676
QY 1597 AGCCACCTCAATGACCTTGTCTGCCAGCTTTACCACTTTGGGCAATTCAGCTCTCACCGTAT 1656
Db 3677 AGCCACCTCAATGACCTTGTCTGCCAGCTTTACCACTTTGGGCAATTCAGCTCTCACCGTAT 3736
QY 1657 AACCTCAGCACTCTCGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCGTCCGAG 1716
Db 3737 AACCTCAGCACTCTCGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCGTCCGAG 3796
QY 1717 GACCGAGTCAAGCAGCTGTCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTT 1776
Db 3797 GACCGAGTCAAGCAGCTGTCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTT 3856
QY 1777 CTTTCCAGCTCTGCTCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTAC 1836
Db 3857 CTTTCCAGCTCTGCTCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTAC 3916
QY 1837 TATATCAACACAGAGACTCAAAACAACCTTGTGGAGACCATCCCAAAATGACAGAGCTTAC 1896
Db 3917 TATATCAACACAGAGACTCAAAACAACCTTGTGGAGACCATCCCAAAATGACAGAGCTTAC 3976
QY 1897 CAGCTTTAGCTGACCTGAAATAATGTCCAGATTTCTCAGCTTATAGGACTGCCATGAACTC 1956
Db 3977 CAGCTTTAGCTGACCTGAAATAATGTCCAGATTTCTCAGCTTATAGGACTGCCATGAACTC 4036
QY 1957 CGAAGACTGCGAGAGGCGCTTGTCTGGATCTCTTGAGCCTGTGAGCTGATGTGATGCC 2016
Db 4037 CGAAGACTGCGAGAGGCGCTTGTCTGGATCTCTTGAGCCTGTGAGCTGATGTGATGCC 4096
QY 2017 TTGACACAGCACACACCTCAAGCAAAATGACAGCCCATGATATCTGCAAGATTTAAT 2076
Db 4097 TTGACACAGCACACACCTCAAGCAAAATGACAGCCCATGATATCTGCAAGATTTAAT 4156
QY 2077 TGTGACCACTATTTATGACCCCTGGAGCAAGCAGACAAATTTGGTCAAGCTCCCT 2136
Db 4157 TGTGACCACTATTTATGACCCCTGGAGCAAGCAGACAAATTTGGTCAAGCTCCCT 4216
QY 2137 CTCTGCTGATATGTCTGAACTGGCTCTGAAATGTTTATGATAGCGGACGAAAGGG 2196
Db 4217 CTCTGCTGATATGTCTGAACTGGCTCTGAAATGTTTATGATAGCGGACGAAAGGG 4276
QY 2197 AGGATCCGCTGCTCTTTTAAAACCTGGCATCATTTCCCTGTGTAAGCAGCATTTGGAA 2256
Db 4277 AGGATCCGCTGCTCTTTTAAAACCTGGCATCATTTCCCTGTGTAAGCAGCATTTGGAA 4336
QY 2257 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACCGCGC 2316
Db 4337 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACCGCGC 4396
QY 2317 AGGCTGGCGCTCTCTGCAATTTCTATCCAAATTCAGAGAGAGAGAGAGAGAGAGAGAG 2376
Db 4397 AGGCTGGCGCTCTCTGCAATTTCTATCCAAATTCAGAGAGAGAGAGAGAGAGAGAGAG 4456
QY 2377 TCCCTTTGGGGGAGTAACTTTGAGC 2401
Db 4457 TCCCTTTGGGGGAGTAACTTTGAGC 4481

RESULT 10

ABK81999

ID ABK81999 standard; DNA; 5462 BP.

XX ABK81999;

AC AC

DT 13-AUG-2002 (first entry)

XX XX

DE DNA encoding mini-dystrophin protein deltaR2-R21+H3.

XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
KW

OS	Homo sapiens.
OS	Synthetic.

XX
FN
WO200229056-A2.

11-APR-2002.

04-OCT-2001: 2001WO-US311126.

06-OCT-2000: 2000US-238848P.

PA (UNMT) UNTV MICHIGAN.

XX
PI
Chamberlain JS. Harper SO:

XX
DR
WPT: 2002-435334/46

xx A composition for preparing therapeutic drugs, has a mini-dystrophin PT peptide comprising a specific number of spectrin-like repeat domains, PT or a nucleic acid sequence encoding the mini-dystrophin peptide - PT

PS Disclosure: Fig 14: 145pp: English.

The invention describes a composition comprising a mini-dystrophin peptide comprising a spectrin-like repeat domain, where the domain comprises *n* spectrin-like repeats, and contains no more than *n* spectrin-like repeats, where *n* is an even number between 4-24, or a nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin peptide or the polynucleotide encoding it is useful as a medicament for preparing a drug for therapeutic application and in the preparation of a composition for treatment of muscle disease, e.g. Duchenne's muscular dystrophy (DMD). This sequence represents a mini-dystrophin sequence of the invention.

Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 other;

Query Match 76.6%: Score 1839.4; DB 24; Length 5462;

Best Local Similarity 88.1%; Pred. No. 0;

Matches 2115; Conservative	0; Mismatches	76; Indels	210; Gaps	2;
----------------------------	---------------	------------	-----------	----

Qy 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 60

Db
1199 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 1258

61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCACAAAGGAGAGATTCTCTAAT 120

db
1259 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCATAAT 1318

121 GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATGGATTGACA 180

1319 GATGTGGAAGTGGTGAAAGACCAGTTTCATACTCATGAGGGGTACATGATCGATTTGACA 1378

181 GCCCATCAGGGCCGGTGGTAATA TTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGA 240

Dp 1379 GCCCATCAGGGCCGGGTGGTAATA TTCTACAA TTGGGAAGTAAGCTGATGGGAACAGGA 1438

241 AATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGA 300

1439 A A T T A T C A G A G A T G A A G A A C T G A A G T A C A G A G C A G A T G A A T C T C T A A A T T C A A G A 1498

301 TGGGAAATGCCCTCAGGGTAGGCTAGCATGGAAAAACAAGCAATTACATAGAGTTTAAAG 360

1498: TGGGATCCCTCAGCGTACCTAGCATGGGAAACAACAGCAATTTACATGCTCTGGACTG 1558

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84


```
2489 CAGCCGCTGGGCGATCTCCTCATCTGACTCTCTCCAAAGATCATCTCGAGAAAGTCAAGGCA 2548
1561 CTTGAGGAGAAATTCGCTCTGAAGAGAGACGTCGACCGACGTCATGACCTTGTCTGCG 1620
2549 CTTGAGGAGAAATTCGCTCTGAAGAGAGACGTCGACCGACGTCATGACCTTGTCTGCG 2608
1621 CAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTCGAAGACCTG 1680
2609 CAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTCTCGAAGACCTG 2668
1681 PACACAGATGGAAGCTTCTGAGGTGCGCTCGAGGACCGAGTCAGCAGCACTGCATGAA 1740
2669 AACACAGATGGAAGCTTCTGAGGTGCGCTCGAGGACCGAGTCAGCAGCACTGCATGAA 2728
1741 GCCACAGGCACTTTGGTCCAGCACTCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCC 1800
2729 GCCACAGGCACTTTGGTCCAGCACTCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCC 2788
1801 TGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAGCACTCAACA 1860
2789 TGGGAGAGGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCAGCACTCAACA 2848
1861 ACTTGTCTGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATAAT 1920
2849 ACTTGTCTGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATAAT 2908
1921 GTCAGATCTCAGCTTATAGGACTGCCATGAACTCCGAGACTCGAGAGCGGCTTTGC 1980
2909 GTCAGATCTCAGCTTATAGGACTGCCATGAACTCCGAGACTCGAGAGCGGCTTTGC 2968
1981 TTGGATCTTTGAGCTCTCAGCTGCATGTGATGCTTGGACGACCACTCAAGCAA 2040
2969 TTGGATCTTTGAGCTCTCAGCTGCATGTGATGCTTGGACGACCACTCAAGCAA 3028
2041 AATGACCAAGCCATGATATCTCTGAGATTTATTAATTTGTTGACCACTATTATGACCGC 2100
3029 AATGACCAAGCCATGATATCTCTGAGATTTATTAATTTGTTGACCACTATTATGACCGC 3088
2101 CTGGAGCAAGACACAACTTTGTCAGCTCTCTCTGCTGCTGATGCTGCTGAC 2160
3089 CTGGAGCAAGACACAACTTTGTCAGCTCTCTCTGCTGCTGATGCTGCTGAC 3148
2161 TGGCTGCTGAATGTTTATGATACGGGACCAACAGGAGGATCCGCTGCTCTCTTTTAA 2220
3149 TGGCTGCTGAATGTTTATGATACGGGACCAACAGGAGGATCCGCTGCTCTCTTTTAA 3208
2221 ACTGGCATCATTTCCCTGTGTAAGACACATTTGGAGACAAGTACAGATACCTTTTCAAG 2280
3209 ACTGGCATCATTTCCCTGTGTAAGACACATTTGGAGACAAGTACAGATACCTTTTCAAG 3268
2281 CAAGTGGCAAGTTCACAGGATTTTGTACCGAGCGAGGCTGGGCTCTCTCTCATGAT 2340
3269 CAAGTGGCAAGTTCACAGGATTTTGTACCGAGCGAGGCTGGGCTCTCTCTCATGAT 3328
2341 TCTATCCAAATTCGAAGCAGTTGGGTGAAGTGTGATCCTTTGGGGGCGAGTAACATTGAG 2400
3329 TCTATCCAAATTCGAAGCAGTTGGGTGAAGTGTGATCCTTTGGGGGCGAGTAACATTGAG 3388
2401 C 2401
3389 C 3389
```

RESULT 11

AD37238

D AAD37238 standard; DNA; 3531 BP.

IC AAD37238;

XX 21-AUG-2002 (first entry)

XX Human dystrophin minigene delta3531.

```
Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.
Homo sapiens.
WO200183695-A2.
08-NOV-2001.
27-APR-2001; 2001WO-US13677.
28-APR-2000; 2000US-200777P.
(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
```

New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -

Example 1; Page 50-51; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus, hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 other;

Query Match	72.3%;	Score 1737;	DB 24;	Length 3531;
Best Local Similarity	96.7%;	Pred. No. 0;		
Matches 1765;	Conservative	0;	Mismatches 61;	Indels 0;
Gaps	0;			
Qy	576	AGATCAGCAACTGCTGCTTTGGAGAACAACTTAAGGTATTGGGAGATCGATGGCAAA	635	
Db	1248	AGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGGAATG	1307	
Qy	636	CATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAGACACTCATAGATTACTGCA	695	
Db	1308	CTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAACTCATAGATTACTGCA	1367	
Qy	696	ACAGTTCCCTCCCTGGACCTGGAAAAAGTTTCTTCGCTGGCTTACAGAGCTGAAACACTGC	755	
Db	1368	ACAGTTCCCTCCCTGGACCTGGAAAAAGTTTCTTCGCTGGCTTACAGAGCTGAAACACTGC	1427	
Qy	756	CAATGCTCTACAGGATGCTACCCGTTAGGAAAGGCTCTTAGAGACTCCAGGGAGTAA	815	
Db	1428	CAATGCTCTACAGGATGCTACCCGTTAGGAAAGGCTCTTAGAGACTCCAGGGAGTAA	1487	
Qy	816	AGAGCTGATGAACCAATGCAAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTTA	875	
Db	1488	AGAGCTGATGAACCAATGCAAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTTA	1547	
Qy	876	TCACAACTGGATGAAAAACAGCCAAAAATCTGAGATCCCTGGAGGTTCCGATGATGC	935	
Db	1548	TCACAACTGGATGAAAAACAGCCAAAAATCTGAGATCCCTGGAGGTTCCGATGATGC	1607	
Qy	936	AGTCTCTTACAAAGACGTTTGGATTAACATGAATCTCAAGTGGAGTGAATTCGGAAAAA	995	

Db 1608 AGTCCTGTTACAAAGAGCGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTCGGAAAAA 1667
Qy 996 GTCTCTCAACATTAGTCTCCCATTTGGAAGCCAGTCTCTGACCACTGAGAGCGCTCTGCACCT 1055
Db 1668 GTCTCTCAACATTAGTCTCCCATTTGGAAGCCAGTCTCTGACCACTGAGAGCGCTCTGCACCT 1727
Qy 1056 TTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCC 1115
Db 1728 TTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCC 1787
Qy 1116 ACCTATTTGAGGGGACCTTTCCAGCAGTCTCAGAAGCAGACGATGTACATAGGGCCCTTCAA 1175
Db 1788 ACCTATTTGAGGGGACCTTTCCAGCAGTCTCAGAAGCAGACGATGTACATAGGGCCCTTCAA 1847
Qy 1176 GAGGGAATTGAAAACTAAAGAACTCTAATCATGAGTACTCTTTGAGACTGTACGAAATATT 1235
Db 1848 GAGGGAATTGAAAACTAAAGAACTCTAATCATGAGTACTCTTTGAGACTGTACGAAATATT 1907
Qy 1236 TCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACAGAGGAGCCGAGAGACTGCC 1295
Db 1908 TCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACTCTACAGAGGAGCCGAGAGACTGCC 1967
Qy 1296 TCTGAGGAGAGAGCCAGAACTGTCTACTCGGCTTCTACGAAAGCAGGCTGAGAGGCTCAA 1355
Db 1968 TCTGAGGAGAGAGCCAGAACTGTCTACTCGGCTTCTACGAAAGCAGGCTGAGAGGCTCAA 2027
Qy 1356 TACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTCGCAGAGAGAAAAATAGATGAGAC 1415
Db 2028 TACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTCGCAGAGAGAAAAATAGATGAGAC 2087
Qy 1416 CTTGAAAGACTCCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTCGGCCA 1475
Db 2088 CTTGAAAGACTCCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTCAAGCTCGGCCA 2147
Qy 1476 AGCTGAGTGGATCAAGGGATCTCTGGCAGCCCGTGGCGGATCTCTCTCACTCTCTCCA 1535
Db 2148 AGCTGAGTGGATCAAGGGATCTCTGGCAGCCCGTGGCGGATCTCTCTCACTCTCTCCA 2207
Qy 1536 AGATCAGCTCGAGAAAGTCAAGCACTTCGAGGAGAAATGGCCCTCTGAAAGAGAAAGT 1595
Db 2208 AGATCAGCTCGAGAAAGTCAAGCACTTCGAGGAGAAATGGCCCTCTGAAAGAGAAAGT 2267
Qy 1596 GAGCCAGCTCAATGACCTTGCTCGCAGCTTACCACTTTGGCGATTCAGCTCTCACCCGTA 1655
Db 2268 GAGCCAGCTCAATGACCTTGCTCGCAGCTTACCACTTTGGCGATTCAGCTCTCACCCGTA 2327
Qy 1656 TAACCTCAGCACTCTGGAAGACTGAACACCAAGATGGAAGCTTCTGCAAGTGGCGCTGGA 1715
Db 2328 TAACCTCAGCACTCTGGAAGACTGAACACCAAGATGGAAGCTTCTGCAAGTGGCGCTGGA 2387
Qy 1716 GGACCGAGTCAGGAGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTT 1775
Db 2388 GGACCGAGTCAGGAGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTT 2447
Qy 1776 TCTTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTA 1835
Db 2448 TCTTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTA 2507
Qy 1836 CTATATCAACCAAGAGACTCAACAACTTCTGGGAGCCATCCCAAAATGACAGACTCTA 1895
Db 2508 CTATATCAACCAAGAGACTCAACAACTTCTGGGAGCCATCCCAAAATGACAGACTCTA 2567
Qy 1896 CCAGTCTTTAGCTGACCTGAATATGTCAATTTCTCAGCTTTATAGACTTGCATGAAACT 1955
Db 2568 CCAGTCTTTAGCTGACCTGAATATGTCAATTTCTCAGCTTTATAGACTTGCATGAAACT 2627
Qy 1956 CGAAGACTCGAAGGCCCTTTGCTTGTGATCTCTTGGAGCCCTGTGAGCTGATGTATGATGC 2015
Db 2628 CGAAGACTCGAAGGCCCTTTGCTTGTGATCTCTTGGAGCCCTGTGAGCTGATGTATGATGC 2687
Qy 2016 CTTGGACACACAACTCAAGCAAAATGACAGCCCATGGATATCTCTGCAATATTAA 2075
Db 2688 CTTGGACACACAACTCAAGCAAAATGACAGCCCATGGATATCTCTGCAATATTAA 2747

Qy 2076 TTGTTTGACCACTATTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAACGTCCC 2135
Db 2748 TTGTTTGACCACTATTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAACGTCCC 2807
Qy 2136 TCTCTGCGTGGATATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGG 2195
Db 2808 TCTCTGCGTGGATATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGG 2867
Qy 2196 GAGATCCGCTGCTCTCTTTTAAACTGTCATCATTTCCCTGTGTAAAGACACATTGGA 2255
Db 2868 GAGATCCGCTGCTCTCTTTTAAACTGTCATCATTTCCCTGTGTAAAGACACATTGGA 2927
Qy 2256 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACACAGG 2315
Db 2928 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTACACAGG 2987
Qy 2316 CAGCTGGGCTCTCTCTGATGATTTCTCAAAATTCCTGATGATTTCAAGACAGTTGGTGAAGTTGC 2375
Db 2988 CAGCTGGGCTCTCTCTGATGATTTCTCAAAATTCCTGATGATTTCAAGACAGTTGGTGAAGTTGC 3047
Qy 2376 ATCTTTGGGGGAGTAACATTGAGC 2401
Db 3048 ATCTTTGGGGGAGTAACATTGAGC 3073

RESULT 12

AAD37258
ID AAD37258 standard; DNA; 4498 BP.
XX AC AAD37258;
XX DT 21-AUG-2002 (first entry)
XX Adeno-associated virus vector plasmid, AAV-MCK-3531.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; de.
XX OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX WO200183695-A2.
XX PN 08-NOV-2001.
PD 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
PR (XIAO/) XIAO X.
XX XIAO X;
PI WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX Example 1; Page 62-63; 7ipp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.

XX
SQ Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 other;

Query Match 72.3%; Score 1737; DB 24; Length 4498;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 576 AGATCAGCGCAACTGCTGCTTTGGAAGCAACAACTTAAGGTATTGGGAGATCGATGGGCAAA 635
DB 2005 AGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATG 2064

QY 636 CATCTGTAGATGACAGAGAACCGCTGGTCTTTTACAGACACTCATAGATTACTGCA 695

DB 2065 CCTCAGGGTAGCTAGCATGTGAAAACAAAGCAATTTACATAGAACTCATAGATTACTGCA 2124

QY 696 ACAGTTCCCTCGACCTCGAAAAGTTTCTTGCTGCTTACAGAACTGAAACAACTGC 755

DB 2125 ACAGTTCCCTCGACCTGAAAAGTTTCTTGCTGCTTACAGAACTGAAACAACTGC 2184

QY 756 CAATGCTCTACAGATGCTACCGTAAAGAAAGGCTCTTAGAAGACTCCAAGGGATAAA 815

DB 2185 CAATGCTCTACAGATGCTACCGTAAAGAAAGGCTCTTAGAAGACTCCAAGGGATAAA 2244

QY 816 AGAGCTGATGAACAATGSCAACGACTCCAAGGTGAATTTGAAGCTCACACAGATGTTTA 875

DB 2245 AGAGCTGATGAACAATGSCAACGACTCCAAGGTGAATTTGAAGCTCACACAGATGTTTA 2304

QY 876 TCACAACTGGATGAACAAGCAAGCAAAATTCCTGAGATCCCTGGAAAGTTCCGATGATGC 935

DB 2305 TCACAACTGGATGAACAAGCAAGCAAAATTCCTGAGATCCCTGGAAAGTTCCGATGATGC 2364

QY 936 AGTCCGTGTTACAAGAGCTTTGATATACATGAATTTCAAGTGGAGTGAATTTCCGAAATA 995

DB 2365 AGTCCGTGTTACAAGAGCTTTGATATACATGAATTTCAAGTGGAGTGAATTTCCGAAATA 2424

QY 996 GTCTCTCAACATTAGTCCCATTTGGAAGCCAGTCTGACCACTGGAAGCGTCTGCACCT 1055

DB 2425 GTCTCTCAACATTAGTCCCATTTGGAAGCCAGTCTGACCACTGGAAGCGTCTGCACCT 2484

QY 1056 TTCTCTCGAGAACTTCTGTGTGGCTACAGCTGGAAGATGATGAATTAAGCCGCGAGGC 1115

DB 2485 TTCTCTCGAGAACTTCTGTGTGGCTACAGCTGGAAGATGATGAATTAAGCCGCGAGGC 2544

QY 1116 ACCTATTGGAGGCGACTTCCAGCAGTTGAGAGCAGACGATGTACATAGGCGCTTCAA 1175

DB 2545 ACCTATTGGAGGCGACTTCCAGCAGTTGAGAGCAGACGATGTACATAGGCGCTTCAA 2604

QY 1176 GAGGGAATTGAAACCTTAAGAACCTGTATCATGATGATCTTTGAGACTGTACGAATATT 1235

DB 2605 GAGGGAATTGAAACCTTAAGAACCTGTATCATGATGATCTTTGAGACTGTACGAATATT 2664

QY 1236 TCTGACAGAGCGCTTTGGAAGGCTAGAGAACTCTACAGGAGCCAGAGAGCTGCC 1295

DB 2665 TCTGACAGAGCGCTTTGGAAGGCTAGAGAACTCTACAGGAGCCAGAGAGCTGCC 2724

QY 1296 TCCTGAGGAGAGCGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAA 1355

DB 2725 TCCTGAGGAGAGCGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAA 2784

QY 1356 TACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 1415

DB 2785 TACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGAC 2844

QY 1416 CTTTGAAGAGCTCCAGGAATTTCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGGCCA 1475

DB 2845 CTTTGAAGAGCTCCAGGAATTTCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGGCCA 2904

QY 1476 AGCTGAGGTGATCAAGGGATTCCTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCA 1535

DB 2905 AGCTGAGGTGATCAAGGGATTCCTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCA 2964

RESULT 13

ID AAD06794 standard; DNA; 5952 BP.

AC AAD06794;

XX 06-AUG-2001 (first entry)

DE Human dystrophin gene (Becker form).

XX Human; dystrophin; extein; intein; trans-splicing; gene therapy;

QY 1536 AGATCACTCGAGAAAGTCAAGGCATCTCGAGGAGAAATTTGGCTCTGAAAGAGAACGT 1595

DB 2965 AGATCACTCGAGAAAGTCAAGGCATCTCGAGGAGAAATTTGGCTCTGAAAGAGAACGT 3024

QY 1596 GAGCAGGTCAATGACCTTGTCTGCCAGCTTACACATTTGGGCAATTCAGCTCTCACCGTA 1655

DB 3025 GAGCAGGTCAATGACCTTGTCTGCCAGCTTACACATTTGGGCAATTCAGCTCTCACCGTA 3084

QY 1656 TAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGA 1715

DB 3085 TAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGA 3144

QY 1716 GGACCGAGTCAGGAGCTGCATGAGGCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 1775

DB 3145 GGACCGAGTCAGGAGCTGCATGAGGCCACAGGGACTTTGGTCCAGCATCTCAGCACTT 3204

QY 1776 TCTTTTCAACGCTCTCTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTA 1835

DB 3205 TCTTTTCAACGCTCTCTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTA 3264

QY 1836 CTATATCAACCAACAGACTCAAACTTCTGCTGGGACCATCCCAAAATGACAGCTCTTA 1895

DB 3265 CTATATCAACCAACAGACTCAAACTTCTGCTGGGACCATCCCAAAATGACAGCTCTTA 3324

QY 1896 CCAGTCTTTAGCTGACCTGAATATGTCAGATTTCTAGCTTATAGGACTGCCATGAACT 1955

DB 3325 CCAGTCTTTAGCTGACCTGAATATGTCAGATTTCTAGCTTATAGGACTGCCATGAACT 3384

QY 1956 CCGAAGACTCGAAGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGCATGTGATGC 2015

DB 3385 CCGAAGACTCGAAGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGCATGTGATGC 3444

QY 2016 CTTGGAACGACCAACCTCAAGCAAAATGACCCCATCGATATCTCTGAGATTTTAA 2075

DB 3445 CTTGGAACGACCAACCTCAAGCAAAATGACCCCATCGATATCTCTGAGATTTTAA 3504

QY 2076 TTGTTTTCACCACTATTATGACCCCTTGGAGCAAGACCAACAAATTTGCTCAACGTCCT 2135

DB 3505 TTGTTTTCACCACTATTATGACCCCTTGGAGCAAGACCAACAAATTTGCTCAACGTCCT 3564

QY 2136 TCTCTGCTGATATGCTGAACTGCTGCTGAATGTTTATGATACGGGACGAACAGG 2195

DB 3565 TCTCTGCTGATATGCTGAACTGCTGCTGAATGTTTATGATACGGGACGAACAGG 3624

QY 2196 GAGGATCCGTGCTGCTTTTAAACTGGCATCATTTCCCTGCTGTAAGCAATTTGGA 2255

DB 3625 GAGGATCCGTGCTGCTTTTAAACTGGCATCATTTCCCTGCTGTAAGCAATTTGGA 3684

QY 2256 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGG 2315

DB 3685 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGG 3744

QY 2316 CAGGCTGGGCTCTCTTCTGATGATTTCTATCCAATTCGAAGACAGTTGGTGAAGTTGC 2375

DB 3745 CAGGCTGGGCTCTCTTCTGATGATTTCTATCCAATTCGAAGACAGTTGGTGAAGTTGC 3804

QY 2376 ATCCTTTTGGGGGCGAGTAACTTGGAGC 2401

DB 3805 ATCCTTTTGGGGGCGAGTAACTTGGAGC 3830

Duchenne muscular dystrophy; Becker muscular dystrophy; DMD; BMD; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT misc_recomb

FT 2847..2848

FT /tag= a

FT /label= S4 junction site

FT /note= "Dystrophin Gene is split at this site and

FT attached to an intron sequence. The resulting

FT fragment is used to produce plasmid PSD4"

FT 2952..2953

FT /tag= b

FT /label= S3 junction site

FT /note= "Dystrophin Gene is split at this site and

FT attached to an intron sequence. The resulting

FT fragment is used to produce plasmid PSD3"

FT 3198..3199

FT /tag= c

FT /label= S2 junction site

FT /note= "Dystrophin Gene is split at this site and

FT attached to an intron sequence. The resulting

FT fragment is used to produce plasmid PSD2"

FT 3300..3301

FT /tag= d

FT /label= S1 junction site

FT /note= "Dystrophin Gene is split at this site and

FT attached to an intron sequence. The resulting

FT fragment is used to produce plasmid PSD1"

XX W0200129243-A1.

XX 26-APR-2001.

XX 13-OCT-2000; 2000WO-CA01216.

XX 15-OCT-1999; 99US-0159868.

XX (UYDA-) UNIV DALHOUSIE.

XX (UYPI-) UNIV PITTSBURGH.

XX Paul XL, Xiao X;

XX WPI; 2001-367297/38.

XX Use of spontaneous or automatic protein splicing to join two or more
XX peptides at junction site involves expressing extein peptides having
XX co-reacting portions of split intein attached to them, so that peptides
XX splice .

XX Example 1; Fig 2; 81pp; English.

XX The invention relates to a method directed to the use of spontaneous
XX or automatic protein trans-splicing to join two or more peptides at
XX junction site. This method involves expressing extein peptides having
XX co-reacting portions of split intein attached to them, so that peptides
XX will splice automatically under suitable conditions. The invention also
XX provides methods for circumventing virion packaging size limitations in
XX recombinant virus particles, by splitting a coding region for a protein
XX to be delivered into two or more extein genes, which are packaged in
XX separate virus particles and are co-delivered in a target cell for the
XX expression and for subsequent trans-splicing to form the complete
XX protein. In particular, the method is used for trans-splicing human
XX dystrophin and in gene therapies of recombinant adeno-associated virus
XX (AAV) particles that encode trans-spliced dystrophin, for treating
XX diseases such as Duchenne muscular dystrophy (DMD) or Becker muscular
XX dystrophy (BMD). The present sequence is human dystrophin gene
XX (Becker form).

XX Sequence 5952 BP; 1860 A; 1344 C; 1410 G; 1338 T; 0 other;

XX Query Match 72.3%; Score 1735.6; DB 22; Length 5952;

XX Best Local Similarity 98.8%; Pred. No. 0;

		Matches 1750;	Conservative	0;	Mismatches	24;	Indels	0;	Gaps	0;
QY	628	TGGGCAAA	CATCTGTAGATGGACAGACCGCTGGTCTTTTACAAGACACTCATAGA	687						
DB	2902	TGAGAAAGCA	TTTCAAAAGGGTGAAGGAGAGGCTGCTTTTGAAGAACTCATAGA	2961						
QY	688	TTACTGCAAC	AGTTCCCTGGACCTGGAAAAGTTTCTTGCTTGCTTACAGAAGCTGAA	747						
DB	2962	TTACTGCAAC	AGTTCCCTGGACCTGGAAAAGTTTCTTGCTTGCTTACAGAAGCTGAA	3021						
QY	748	ACAACCTGCA	ATGCTCTACAGGATGCTACCCGTAAAGGAAGCTCCTTAGAGACTCCAAG	807						
DB	3022	ACAACCTGCA	ATGCTCTACAGGATGCTACCCGTAAAGGAAGCTCCTTAGAGACTCCAAG	3081						
QY	808	GGAGTAAAGA	AGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA	867						
DB	3082	GGAGTAAAGA	AGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA	3141						
QY	868	GATGTTTATC	AACTGGATGAACAAGCCAAATAATCCTGAGATCCTGGAGGTTCC	927						
DB	3142	GATGTTTATC	AACTGGATGAACAAGCCAAATAATCCTGAGATCCTGGAGGTTCC	3201						
QY	928	GATGATGCA	GTCTCTGTTACAAAGAGCTTTTGGATAACATGAACCTTCAAGTGGAGTGAATT	987						
DB	3202	GATGATGCA	GTCTCTGTTACAAAGAGCTTTTGGATAACATGAACCTTCAAGTGGAGTGAATT	3261						
QY	988	CGGAAAAGCT	CTCTCAACATTTAGTCCCATTTGGAGCCAGTTCTGACAGTGGAGCGT	1047						
DB	3262	CGGAAAAGCT	CTCTCAACATTTAGTCCCATTTGGAGCCAGTTCTGACAGTGGAGCGT	3321						
QY	1048	CTGCACCTTT	CTCTCGAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC	1107						
DB	3322	CTGCACCTTT	CTCTCGAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC	3381						
QY	1108	CGCAGGCACT	TATTGGAGCGCACTTTCCAGCAGTTTTCAGAGCAGCAAGATGATGATAGG	1167						
DB	3382	CGCAGGCACT	TATTGGAGCGCACTTTCCAGCAGTTTTCAGAGCAGCAAGATGATGATAGG	3441						
QY	1168	GCCTTCAAG	AGGGAAATTGAAACTTAAGAACTGTAATCATGAGTACTCTTGAGACTGTA	1227						
DB	3442	GCCTTCAAG	AGGGAAATTGAAACTTAAGAACTGTAATCATGAGTACTCTTGAGACTGTA	3501						
QY	1228	CGAATATTT	CTGACAGAGCAGCTTTTGGAAAGCACTAGAGAACTCTACAGAGGCCAGA	1287						
DB	3502	CGAATATTT	CTGACAGAGCAGCTTTTGGAAAGCACTAGAGAACTCTACAGAGGCCAGA	3561						
QY	1288	GAGTCGCTCT	TGAGAGAGAGCCAGATGTCATCGGCTTCTACGAAAGCAGGCTGAG	1347						
DB	3562	GAGTCGCTCT	TGAGAGAGAGCCAGATGTCATCGGCTTCTACGAAAGCAGGCTGAG	3621						
QY	1348	GAGGTCAATA	CTGAGTGGGAAAATAATTGAACCTGCACTCCGCTGACTGGCAGAAAAATA	1407						
DB	3622	GAGGTCAATA	CTGAGTGGGAAAATAATTGAACCTGCACTCCGCTGACTGGCAGAAAAATA	3681						
QY	1408	GATGAGACCT	TGAAAGATCCAGGAATTCAGAGGCCACGGATGAGCTGACCTCAAG	1467						
DB	3682	GATGAGACCT	TGAAAGATCCAGGAATTCAGAGGCCACGGATGAGCTGACCTCAAG	3741						
QY	1468	CTCGGCCAAG	CTGAGTGATCAAGGGATCCTGGCAGCCGCTGGGCGATCTCTCTCAATTGAC	1527						
DB	3742	CTCGGCCAAG	CTGAGTGATCAAGGGATCCTGGCAGCCGCTGGGCGATCTCTCTCAATTGAC	3801						
QY	1528	TTCTCTCAAG	ATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGGCTCTGAAA	1587						
DB	3802	TTCTCTCAAG	ATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGGCTCTGAAA	3861						
QY	1588	GAGAAGCTGA	GGCCACCTCAATGACCTTGTCTGCGCAGCTTACCACCTTTGGGCATTCAGCTC	1647						
DB	3862	GAGAAGCTGA	GGCCACCTCAATGACCTTGTCTGCGCAGCTTACCACCTTTGGGCATTCAGCTC	3921						
QY	1648	TCACCGTATA	AACTCTAGCAGCTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTTG	1707						
DB	3922	TCACCGTATA	AACTCTAGCAGCTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTTG	3981						

QY 1708 GCCGTGAGGACCGAGTCAGCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCT 1767
Db |||||
QY 3982 GCCGTGAGGACCGAGTCAGCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCT 4041
Db |||||
QY 1768 CAGCAGCTTTCTTCCAGCTGTCTCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAA 1827
Db |||||
QY 4042 CAGCAGCTTTCTTCCAGCTGTCTCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAA 4101
Db |||||
QY 1828 GTGCCCTACTATATCAACACGAGACTCAACAACTTGTCTGGGACCATCCCAAAATGACA 1887
Db |||||
QY 4102 GTGCCCTACTATATCAACACGAGACTCAACAACTTGTCTGGGACCATCCCAAAATGACA 4161
Db |||||
QY 1888 GAGCTCTACCAAGCTTTTGTCTGACCTGAATATGTTCAGATTTCTCAGCTTATAGGACTGCC 1947
Db |||||
QY 4162 GAGCTCTACCAAGCTTTTGTCTGACCTGAATATGTTCAGATTTCTCAGCTTATAGGACTGCC 4221
Db |||||
QY 1948 ATGAAGCTCCGAGACTGCGAGAGCCCTTTGTCTGGATCTCTTGAGCTGTGAGCTGCA 2007
Db |||||
QY 4222 ATGAAGCTCCGAGACTGCGAGAGCCCTTTGTCTGGATCTCTTGAGCTGTGAGCTGCA 4281
Db |||||
QY 2008 TGTGATGCTTGGACACGACCAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAG 2067
Db |||||
QY 4282 TGTGATGCTTGGACACGACCAACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAG 4341
Db |||||
QY 2068 ATATTATTTGTTGACCACTATTATATGACCGCTGGAGCAAGACACAAATTTGGTC 2127
Db |||||
QY 4342 ATATTATTTGTTGACCACTATTATATGACCGCTGGAGCAAGACACAAATTTGGTC 4401
Db |||||
QY 2128 AACGTCCCTCTCTCGTGGATATGTCTGAACCTGGCTGTGAATGTTTATGATACGGGA 2187
Db |||||
QY 4402 AACGTCCCTCTCTCGTGGATATGTCTGAACCTGGCTGTGAATGTTTATGATACGGGA 4461
Db |||||
QY 2188 CGAAGAGGAGGATCCGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
Db |||||
QY 4462 CGAAGAGGAGGATCCGTCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 4521
Db |||||
QY 2248 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGT 2307
Db |||||
QY 4522 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGT 4581
Db |||||
QY 2308 GACCAGGCGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGT 2367
Db |||||
QY 4582 GACCAGGCGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCGAAGACAGTTGGGT 4641
Db |||||
QY 2368 GAAGTTGATCTTTGGGGGAGTAACATTGAGC 2401
Db |||||
QY 4642 GAAGTTGATCTTTGGGGGAGTAACATTGAGC 4675
Db |||||

RESULT 14

ABR82000
ID ABR82000 standard; DNA; 8689 BP.

AC ABR82000;

DT 13-AUG-2002 (first entry)

DE DNA encoding mini-dystrophin protein deltaH2-R19.

XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.

XX Homo sapiens.
OS Synthetic.

XX WO200229056-A2.

XX 11-APR-2002.

XX 04-OCT-2001; 2001WO-US31126.

XX 06-OCT-2000; 2000US-238848P.

XX (UNMI) UNIV MICHIGAN.
XX Chamberlain JS, Harper SQ;
XX WPI; 2002-435334/46.
XX
XX A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
XX Disclosure; Fig 15; 145pp; English.
XX
XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX
XX Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 other;

Query Match 72.3%; Score 1735.6; DB 24; Length 8689;

Best Local Similarity 98.6%; Pred. No. 0;

Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 628 TGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAGACACTCATAGA 687
Db |||||
QY 2948 TGGAGAGCATTCAATAAGGGTGAAGGAGGCTGCTTTGGAGAACTCATAGA 3007
Db |||||
QY 688 TTACTGCAACAGTTCCTCCCTGCACTCGAAGAAAGTTCTTGCTGGCTTACAGAGCTGAA 747
Db |||||
QY 3008 TTACTGCAACAGTTCCTCCCTGCACTCGAAGAAAGTTCTTGCTGGCTTACAGAGCTGAA 3067
Db |||||
QY 748 ACACTGCAATGCTTACAGAGTCTACCGTAAGGAAGGCTCTTAGAGACTCCCAAG 807
Db |||||
QY 3068 ACACTGCAATGCTTACAGAGTCTACCGTAAGGAAGGCTCTTAGAGACTCCCAAG 3127
Db |||||
QY 808 GGAGTAAAGAGCTGATGAAAACAAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 867
Db |||||
QY 3128 GGAGTAAAGAGCTGATGAAAACAAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 3187
Db |||||
QY 868 GATGTTTATCAACCTGGATGAAAACAGCCAAAATCTTGAGATCCCTGGAGGTTCC 927
Db |||||
QY 3188 GATGTTTATCAACCTGGATGAAAACAGCCAAAATCTTGAGATCCCTGGAGGTTCC 3247
Db |||||
QY 928 GATGATGCAGTCTCTTACAAAGACGTTTGGATTAACATGAACCTTCAAGTGGAGTGAACCTT 987
Db |||||
QY 3248 GATGATGCAGTCTCTTACAAAGACGTTTGGATTAACATGAACCTTCAAGTGGAGTGAACCTT 3307
Db |||||
QY 988 CGGAAAAAGTCTCTCAACATTAGTTCCTCAATTTGGAAGCAGTTCCTGACAGTGAAGCGT 1047
Db |||||
QY 3308 CGGAAAAAGTCTCTCAACATTAGTTCCTCAATTTGGAAGCAGTTCCTGACAGTGAAGCGT 3367
Db |||||
QY 1048 CTGCACCTTTCTCTGCAGGAACCTCTGCTGGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
Db |||||
QY 3368 CTGCACCTTTCTCTGCAGGAACCTCTGCTGGTGGCTACAGCTGAAAGATGATGAATTAAGC 3427
Db |||||
QY 1108 CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCCAGAGCAGAACGATGTTACATAGG 1167
Db |||||
QY 3428 CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCCAGAGCAGAACGATGTTACATAGG 3487
Db |||||
QY 1168 GCCTTCAAGAGGGAATTGAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAG 1227
Db |||||
QY 3488 GCCTTCAAGAGGGAATTGAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAG 3547
Db |||||
QY 1228 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGAGCTAGAGAACTTCTACAGGAGCCGAGA 1287
Db |||||
QY 3548 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGAGCTAGAGAACTTCTACAGGAGCCGAGA 3607
Db |||||

1288 GAGCTGCTCTCTGAGGAGAGCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1347
Db GAGCTGCCCTCTGAGGAGAGAGCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAG 3667
1348 GAGGTCAATCTGAGTGGGAAATTTGAACCTGCACTCGCTGACTGCGCAGAGAAAATA 1407
Db GAGGTCAATCTGAGTGGGAAATTTGAACCTGCACTCGCTGACTGCGCAGAGAAAATA 3727
1408 GATGAGACCTTGAAGACATCCAGGACTTCAAGAGGCCACCGATGAGCTGAGCCTCAAG 1467
Db GATGAGACCTTGAAGACATCCAGGACTTCAAGAGGCCACCGATGAGCTGAGCCTCAAG 3787
1468 CTGCGCAAGCTGAGTGTATCAAGGATCTGCGCAGCCGCTGGCGATCTCTCTATTGAC 1527
Db CTGCGCAAGCTGAGTGTATCAAGGATCTGCGCAGCCGCTGGCGATCTCTCTATTGAC 3847
1528 TCTCTCAAGATCACTCGAGAAAGTCAAGGACTTCAAGGAGAAATTCGCGCTCTGAAA 1587
Db TCTCTCAAGATCACTCGAGAAAGTCAAGGACTTCAAGGAGAAATTCGCGCTCTGAAA 3907
1588 GAGAACTGAGCCACGCTCAATGACCTTGTGCGCAGCTTACCACCTTGGGCACTCAGCTC 1647
Db GAGAACTGAGCCACGCTCAATGACCTTGTGCGCAGCTTACCACCTTGGGCACTCAGCTC 3967
1648 TCACCGTATAAATCTGAGCTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTG 1707
Db TCACCGTATAAATCTGAGCTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTG 4027
1708 GCGCTGAGGAGCCGCTGAGGAGCTGATGAGCCACAGGACTTGTGCTCCAGCATCT 1767
Db GCGCTGAGGAGCCGCTGAGGAGCTGATGAGCCACAGGACTTGTGCTCCAGCATCT 4087
1768 CAGCAGCTTTCTTTCCAGCTCTGTCAGGCTGCTGCGGAGAGAGCCATCTCGCAACAAA 1827
Db CAGCAGCTTTCTTTCCAGCTCTGTCAGGCTGCTGCGGAGAGAGCCATCTCGCAACAAA 4147
1828 GTGCCCTACTATATCAACGAGACTCAACAACTTGTGCGGAGCCATCTCGCAACAAA 1887
Db GTGCCCTACTATATCAACGAGACTCAACAACTTGTGCGGAGCCATCTCGCAACAAA 4207
1888 GAGCTCTACCACTTTTAGTGTGACCTGATATGTGAGATTTCTGAGTTATAGGACTGCC 1947
Db GAGCTCTACCACTTTTAGTGTGACCTGATATGTGAGATTTCTGAGTTATAGGACTGCC 4267
1948 ATGAACTCGAAGACTGCGAAGAGCCCTTGTGAGTCTCTGAGCTGTGAGCTGCA 2007
Db ATGAACTCGAAGACTGCGAAGAGCCCTTGTGAGTCTCTGAGCTGTGAGCTGCA 4327
2008 TGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACAGCCCATGATATCTGCGAG 2067
Db TGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACAGCCCATGATATCTGCGAG 4387
2068 ATTATTAATTGTTTGACCACTATTATGACCGCTGGAGCAAGACAAATTTGCTC 2127
Db ATTATTAATTGTTTGACCACTATTATGACCGCTGGAGCAAGACAAATTTGCTC 4447
2128 AAGTCCCTCTCTGCGTGGATATGTCTGAATGCTGCTGATCTTATGATACGGA 2187
Db AAGTCCCTCTCTGCGTGGATATGTCTGAATGCTGCTGATCTTATGATACGGA 4507
2188 CGAAGAGGAGATCCGCTGCTCTTTTAAACCTGGCATCTTTCCCTGTGTAAGCA 2247
Db CGAAGAGGAGATCCGCTGCTCTTTTAAACCTGGCATCTTTCCCTGTGTAAGCA 4567
2248 CATTGGAAGCAAGTACAGTACCTTTCAAGCAAGTGGCAGTTTCAACAGATTTTGT 2307
Db CATTGGAAGCAAGTACAGTACCTTTCAAGCAAGTGGCAGTTTCAACAGATTTTGT 4627
2308 GACCAGGCGAGCTGGGCTCTCTTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGT 2367
Db GACCAGGCGAGCTGGGCTCTCTTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGT 4687

2368 GAAGTTGCATCTTTGGGGGCGAGTAACATTGAGC 2401
Db GAAGTTGCATCTTTGGGGGCGAGTAACATTGAGC 4721
RESULT 15
AAD37229
ID AAD37229 standard; DNA; 11058 BP.
XX AC AAD37229;
XX DT 21-AUG-2002 (first entry)
XX Human dystrophin protein coding sequence.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX OS Homo sapiens.
XX FN WO200183695-A2.
XX PD 08-NOV-2001.
XX PF 27-APR-2001; 2001WO-US13677.
XX PR 28-APR-2000; 2000US-200777P.
XX PA (XIAO/) XIAO X.
XX PI XIAO X;
XX DR WPI; 2002-049342/06.
XX PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX Example 1; Page 40-43; 71pp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is human
XX dystrophin protein coding sequence.
SQ Sequence 11058 BP; 3686 A; 2292 C; 2621 G; 2459 T; 0 other;
Query Match 72.3%; Score 1735.6; DB 24; Length 11058;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 528 TGGCAACAATCTGTAGATGGACAGAGACCGTGGTCTCTTTTACAGACACTCATAGA 687
Db TGGCAACAATCTGTAGATGGACAGAGACCGTGGTCTCTTTTACAGACACTCATAGA 8067
QY 688 TTACTGCAACAGTTTCCCTGGACCTGGAAAAGTTTCTTCCTCGCTGCTTACAGAAGCTGAA 747
Db TTACTGCAACAGTTTCCCTGGACCTGGAAAAGTTTCTTCCTCGCTGCTTACAGAAGCTGAA 8127
QY 748 ACACTGCAATGTCTTACAGGATGCTACCCGTAAAGAAAGCTCTTAGAGACTCCAAG 807
Db ACACTGCAATGTCTTACAGGATGCTACCCGTAAAGAAAGCTCTTAGAGACTCCAAG 8187
QY 808 GGAGTAAAGAGCTGATGAAACAATGGCAAGACCTCCAGGTGAAATTGAAGCTCACACA 867

Db 8188 GGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAAGGTGAATTTGAAGCTCACACA 8247
2y 868 GATGTTTATCACAACTGATGAACACCCAAATAATCCTGAGTCCCTGGAAGTTCC 927
Db 8248 GATGTTTATCACAACTGATGAACACCCAAATAATCCTGAGATCCCTGGAAGTTCC 8307
2y 928 GATGATGCAGTCCCTGTTTCAAAAGACCTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 987
Db 8308 GATGATGCAGTCCCTGTTTCAAAAGACCTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTT 8367
2y 988 CGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAGGCT 1047
Db 8368 CGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAGGCT 8427
2y 1048 CTCACCTTCTCTGAGGAACTTCTGCTGTGCTACAGCTCAAGATGATCAATTAAGC 1107
Db 8428 CTCACCTTCTCTGAGGAACTTCTGCTGTGCTACAGCTCAAGATGATCAATTAAGC 8487
2y 1108 CGCAGGACCTTATGAGGCGACTTTCAGCAGTTTCAGAGCAACGATGTATCAGG 1167
Db 8488 CGCAGGACCTTATGAGGCGACTTTCAGCAGTTTCAGAGCAACGATGTATCAGG 8547
2y 1168 GCCTTCAAGAGGGAATTTGAAACTTAAGAACCTGTAACTCATGAGTACTCTCAGACTGTA 1227
Db 8548 GCCTTCAAGAGGGAATTTGAAACTTAAGAACCTGTAACTCATGAGTACTCTCAGACTGTA 8607
2y 1228 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGAGCCACA 1287
Db 8608 CGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGAGCCACA 8667
2y 1288 GAGCTGCCTCTGAGAGAGAGCCAGATGTCACCTCGCTTCTACGAAGCAGCTGAG 1347
Db 8668 GAGCTGCCTCTGAGAGAGAGCCAGATGTCACCTCGCTTCTACGAAGCAGCTGAG 8727
2y 1348 GAGCTCAATACTCAGTGGGAAAAATTAACCTCGCTGACTGGCAGAGAAAAATA 1407
Db 8728 GAGGTCAAFACTAGTGGGAAAAATTTGACCTGCACTCCGCTGACTGGCAGAGAAAAATA 8787
2y 1408 GATGAGACCTTTGAAAGACTCCAGAACCTTCAAGAGGCCACGGATGAGCTGACCTCAAG 1467
Db 8788 GATGAGACCTTTGAAAGACTCCAGAACCTTCAAGAGGCCACGGATGAGCTGACCTCAAG 8847
2y 1468 CTGGCCCAAGCTCAGTGTATCAAGGATCTCTGACGCCCTGGCGATCTCTCATTTGAC 1527
Db 8848 CTGGCCCAAGCTCAGTGTATCAAGGATCTCTGACGCCCTGGCGATCTCTCATTTGAC 8907
2y 1528 TCTCTCCAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAA 1587
Db 8908 TCTCTCCAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAA 8967
2y 1588 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTCAGCTC 1647
Db 8968 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTCAGCTC 9027
2y 1648 TCACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAAGATGGAAGCTTCTGAGGCTG 1707
Db 9028 TCACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAAGATGGAAGCTTCTGAGGCTG 9087
2y 1708 GCCGTGAGAACCGAGTCAAGGAGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCT 1767
Db 9088 GCCGTGAGAACCGAGTCAAGGAGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCT 9147
2y 1768 CAGCACCTTTCTTCCAGCTGTCTGCGAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAA 1827
Db 9148 CAGCACCTTTCTTCCAGCTGTCTGCGAGGTCCTCGGAGAGAGCCATCTCGCCAAACAAA 9207
2y 1828 GTGCCCTACTATATCAACACGAGACTCAAAACCTTGTGGAGCCATCCCAAAATGACA 1887
Db 9208 GTGCCCTACTATATCAACACGAGACTCAAAACCTTGTGGAGCCATCCCAAAATGACA 9267
2y 1888 GAGCTCTACAGTCTTTAGTGAACCTGATTAATGTGATGATCTCAGCTTATAGGCTGCC 1947

Db 9268 GAGCTCTACAGTCTTTAGCTGACCTGAATAATGTGATGATCTCAGCTTATAGGACTGCC 9327
2y 1948 ATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTGGATCTCTTGGAGCTCTTGGAGCTGCA 2007
Db 9328 ATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTGGATCTCTTGGAGCTCTTGGAGCTGCA 9387
2y 2008 TGTGATCCCTTGGACCAAGCAACCTCAAGCAAAATGACAGGCCATGGATATCCTGCGAG 2067
Db 9388 TGTGATCCCTTGGACCAAGCAACCTCAAGCAAAATGACAGGCCATGGATATCCTGCGAG 9447
2y 2068 ATTATTAATTTGTTGACCACTATTTATGACCGCTGAGCAAGAGCAACAATTTGGTC 2127
Db 9448 ATTATTAATTTGTTGACCACTATTTATGACCGCTGAGCAAGAGCAACAATTTGGTC 9507
2y 2128 AACGTCCCTCTCTGCGTGGATATGTGTCTGAATCTGCTGAATGTTATGATACGGGA 2187
Db 9508 AACGTCCCTCTCTGCGTGGATATGTGTCTGAATCTGCTGAATGTTATGATACGGGA 9567
2y 2188 CGAACAGGAGGATCCGTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
Db 9568 CGAACAGGAGGATCCGTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 9627
2y 2248 CATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 2307
Db 9628 CATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 9687
2y 2308 GACAGCGCAGGCTGGGCTCTCTTCTGCAATGATCTATCCAAATTTCAAGACAGTTGGGT 2367
Db 9688 GACAGCGCAGGCTGGGCTCTCTTCTGCAATGATCTATCCAAATTTCAAGACAGTTGGGT 9747
2y 2368 GAAGTTGCATCCCTTTGGGGGCGAGTAACTGAGC 2401
Db 9748 GAAGTTGCATCCCTTTGGGGGCGAGTAACTGAGC 9781

Search completed: February 1, 2004, 12:04:30
Job time : 591.303 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:14:55 ; Search time 8553.13 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-9_COPY_1000_3400

Perfect score: 2401
Sequence: 1 gscagtcattgatgagag.....tggggscgtaacattgagc 2401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1839.4	76.6	5462	6	AX538621 Sequence
2	1735.6	72.3	5952	6	AR304538 Sequence
3	1735.6	72.3	5952	6	AX114289 Sequence
4	1735.6	72.3	8689	6	AX538622 Sequence
5	1735.6	72.3	11443	6	AX538624 Sequence
6	1735.6	72.3	12057	6	AX538627 Sequence
7	1735.6	72.3	12446	9	HSMDMR X14298 Human mRNA
8	1735.6	72.3	13957	6	AX409637 Sequence
9	1735.6	72.3	13957	6	AX538581 Sequence
10	1735.6	72.3	13957	9	HUMDYS M18533 Homo sapien
11	1725	71.8	5339	6	AX538620 Sequence
12	1724.6	71.8	13977	6	AR220819 Sequence
13	1540.4	64.2	13887	4	AF070485 Canis fam
14	1506.8	62.8	13815	6	AX306153 Sequence
15	1506.8	62.8	13815	6	AX538582 Sequence
16	1506.8	62.8	13815	10	MUSDYSA M6859 Mouse dysstr
17	1506.8	62.8	19307	6	AR093392 Sequence
18	1506.8	62.8	19307	5	AR142592 Sequence
19	1419	59.1	5417	6	AX538619 Sequence
20	1181.4	49.2	13575	5	GGDYS X13369 Chicken mRN
21	1058.2	44.1	4402	6	E30219 Shortened d
22	859	35.8	4402	6	E30220 Shortened d
23	777.6	32.4	3521	5	AF339031 Danio rer
24	753	31.4	2654	5	FSCDYSTRO M37645 Torpedo cal
25	723.4	30.1	4075	6	E30221 Shortened d
26	719.2	30.0	3747	6	E30218 Shortened d
27	711.2	29.6	10705	10	RNAJ2967 Rattus no
28	708.2	29.5	11096	6	AX538584 Sequence
29	708.2	29.5	11096	10	MMY12229 Y12229 M.musculus
30	690.6	28.8	6045	6	A63605 Sequence 7
31	690.6	28.8	6045	6	AR281528 Sequence
32	690.6	28.8	6059	6	AX107972 Sequence
33	690.6	28.8	10302	6	AX538583 Sequence
34	690.6	28.8	10302	9	HSMQPS X69086 H.sapiens m
35	690.6	28.8	10320	6	A63607 Sequence 9
36	690.6	28.8	10320	6	AR281529 Sequence
37	671.6	28.0	3161	10	MMGTUPH X83506 M.musculus
38	665.6	27.7	3163	6	E30223 Shortened d
39	576.6	24.0	3073	10	AF195788 Rattus no
40	576.6	24.0	3140	10	AF195787 Rattus no
41	576.6	24.0	3499	9	HSU43519 U43519 Human dysstr
42	576.6	24.0	5106	6	AX552248 Sequence
43	565.8	23.6	3275	10	HUMDYS M18025 Mouse dysstr
44	557	23.2	2110	9	HUMDMDX M92650 Human Duch
45	557	23.2	4658	9	EC028720 Homo sapi

ALIGNMENTS

RESULT 1
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
Patent: WO 0229056-A 41 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES		Location/Qualifiers	
source		1..5462	
		/organism="synthetic construct"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:32630"	
		/note="Synthetic"	
BASE COUNT		1668 a 1225 c 1212 g 1357 t	
ORIGIN			
Query Match		76.6%; Score 1839.4; DB 6; Length 5462;	
Best Local Similarity		88.1%; Pred. No. 0;	
Matches 2115; Conservative		0; Mismatches 76; Indels 210; Gaps 2;	
Qy	1	GCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA	60
Db	1199	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA	1258
Qy	61	GTATTATCGTGGCTTCTTCTGCTGAGGACACATGTCRAGACCAAGGAGAGATTCTTAAT	120
Db	1259	GTATTATCGTGGCTTCTTCTGCTGAGGACACATGTCRAGACCAAGGAGAGATTCTTAAT	1318
Qy	121	GATGTGGAAAGTGGTGAAGACACAGATTTCATCTCATGAGGGGTACATGATGATGATTGACA	180
Db	1319	GATGTGAAATGCTGAAGACACAGATTTCATCTCATGAGGGGTACATGATGATGATTGACA	1378
Qy	181	GCCCATCAGGCGCGGTTGGTAAATATCTTACAAATGGGAAGTAAAGCTGATTTGGAAACAGA	240
Db	1379	GCCCATCAGGCGCGGTTGGTAAATATCTTACAAATGGGAAGTAAAGCTGATTTGGAAACAGA	1438
Qy	241	AAATTATCAGAAATGAAGAACTGAAGTACAAAGACAGATCAATCTCTCAATTTCAAGA	300
Db	1439	AAATTATCAGAAATGAAGAACTGAAGTACAAAGACAGATCAATCTCTCAATTTCAAGA	1498
Qy	301	TGGGAATGCTCAGGTAGTACAGTGGAAACAAAGCAATTTACATAGATTTTAATG	360
Db	1499	TGGGAATGCTCAGGTAGTACAGTGGAAACAAAGCAATTTACATAGTCTCTGGAATG	1558
Qy	361	GATCTCCAGATCAGAACTGAAGAGTTGAATGACTGGCTAACAAAACAGAAAGA	420
Db	1559	ACCA-----	1562
Qy	421	ACAAGAAATGGAGGAGAGCCTTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA	480
Db	1563	-----CTATTGGAGCTCTCTCTACTCAGACTGTTACTCTGGTG	1600
Qy	481	CAACACATAGGTGCTTCAAGAGATCTAGACAAAGACAGTCAAGGTCAATTTCTCTC	540
Db	1601	ACAAACCTGTGTTACTAAGGAACTGCCATCTCCAAACTAGAAATGCCATCTTTCCTTG	1660
Qy	541	ACTCACATGTTGTTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA	600
Db	1661	ATGTTGGAG-----	1669
Qy	601	GAACAACTTAAGTATTGGAGATCGATGGCAACATCTGTAGATGGACAGAACCGC	660
Db	1670	-----	1669
Qy	661	TGGGTTCTTTTACAAGACACTCATGATTACTGCAACAGTTCCCTCTGACCTGGAAAG	720
Db	1670	-----CATAGTTACTTGCAACAGTTCCCTCTGACCTGGAAAG	1708
Qy	721	TTTCTTGGCTGTTTACAGAGCTGAAACAACCTGCGAATGTCTTACAGATGCTACCGT	780
Db	1709	TTTCTTGGCTGTTTACAGAGCTGAAACAACCTGCGAATGTCTTACAGATGCTACCGT	1768
Qy	781	AAGGAAGGCTCTAGAGACTCAAGGAGTAAAGAGCTGATGAACAATGGCAAGAC	840
Db	1769	AAGGAAGGCTCTAGAGACTCAAGGAGTAAAGAGCTGATGAACAATGGCAAGAC	1828
Qy	841	CTCCAAGGTGAATTAAGAGCTCACAGATGTTTATCAAACTGGATGAAACAGCCAA	900
Db	1829	CTCCAAGGTGAATTAAGAGCTCACAGATGTTTATCAAACTGGATGAAACAGCCAA	1888
Qy	901	AAAATCCTGAGATCCCTGGAAAGTTCGGATGATGAGTCTCTGTACAAAGAGCTTTGGAT	960
Db	1889	AAAATCCTGAGATCCCTGGAAAGTTCGGATGATGAGTCTCTGTACAAAGAGCTTTGGAT	1948
Qy	961	AACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAAGTCCCATTTG	1020
Db	1949	AACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAAGTCCCATTTG	2008
Qy	1021	GAAGCCAGATTTCTGACCAAGTGGAAAGCTCTGACACTTTCTCTGAGGAACTTTCTGTGTGTG	1080
Db	2009	GAAGCCAGTTCCTGACCAAGTGGAAAGCTCTGACACTTTCTCTGAGGAACTTTCTGTGTGTG	2068
Qy	1081	CTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACCTTATTGGAGCGGACTTTCCAGCA	1140
Db	2069	CTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACCTTATTGGAGCGGACTTTCCAGCA	2128
Qy	1141	GTTTCAGAGCAGAACCATGTACATAGCGGCTTCAAGAGGGAATTTGAAACTAAAGAACCT	1200
Db	2129	GTTTCAGAGCAGAACCATGTACATAGCGGCTTCAAGAGGGAATTTGAAACTAAAGAACCT	2188
Qy	1201	GTAATCATGAGTACTCTTGGAGCTGTACGAATATTTCTGACAGAGCAGCTTTTGAAGGA	1260
Db	2189	GTAATCATGAGTACTCTTGGAGCTGTACGAATATTTCTGACAGAGCAGCTTTTGAAGGA	2248
Qy	1261	CTAGAGAACTCTTACAGAGCCGAGAGCTGCTCTCTGAGGAGAGCCCGAGATGTC	1320
Db	2249	CTAGAGAACTCTTACAGAGCCGAGAGCTGCTCTCTGAGGAGAGCCCGAGATGTC	2308
Qy	1321	ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATPACTGAGTGGGAAAAATTTGAACCTG	1380
Db	2309	ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATPACTGAGTGGGAAAAATTTGAACCTG	2368
Qy	1381	CATCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCAA	1440
Db	2369	CATCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTCAA	2428
Qy	1441	GAGGCCACGATCAGCTGGAGCTCAAGCTCGGCAAGCTGAGGTGATCAAGGATCCTCG	1500
Db	2429	GAGGCCACGATCAGCTGGAGCTCAAGCTCGGCAAGCTGAGGTGATCAAGGATCCTCG	2488
Qy	1501	CAGCCGTTGGCGCATCTCTCATTTGACTCTCTCTCAAGATCACTCGAGAAAGTCAAGGCA	1560
Db	2489	CAGCCGTTGGCGCATCTCTCATTTGACTCTCTCTCAAGATCACTCGAGAAAGTCAAGGCA	2548
Qy	1561	CTTCGAGGAAATTCGGCTCTGAAAGAGAGCTGAGCCAGCTCAATGACCTTGTCTCGC	1620
Db	2549	CTTCGAGGAAATTCGGCTCTGAAAGAGAGCTGAGCCAGCTCAATGACCTTGTCTCGC	2608
Qy	1621	CAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG	1680
Db	2609	CAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTG	2668
Qy	1681	AACACAGATGGAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGAGCTGCATGAA	1740
Db	2669	AACACAGATGGAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGAGCTGCATGAA	2728
Qy	1741	GCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTCTCTCCAGGCTCC	1800
Db	2729	GCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTCTCTCCAGGCTCC	2788
Qy	1801	TGGGAGAGGCCATCTCGGCCAAACAAAGTGGCTTACTATATCAACACAGAGCTCAACA	1860
Db	2789	TGGGAGAGGCCATCTCGGCCAAACAAAGTGGCTTACTATATCAACACAGAGCTCAACA	2848
Qy	1861	ACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTTACCTGACCTGAATAAT	1920
Db	2849	ACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTTACCTGACCTGAATAAT	2908
Qy	1921	GTCAAGTCTCAGCTTATAGGACTGCCATCCGAAACTCCGAAAGACTGCAGAGGCCCTTTTC	1980
Db	2909	GTCAAGTCTCAGCTTATAGGACTGCCATCCGAAACTCCGAAAGACTGCAGAGGCCCTTTTC	2968
Qy	1981	TTGGAATCTTTGAGCCCTGTCAGCTGTCATGTGATGCTTTGGACCGACCAACCTCAAGCAA	2040

Db 2969 TTGATCTCTTGAGCCCTGTCAGCTGATGTGATGCTTGGACACAGACAACTCTAAGCAA 3028
Qy 2041 AATGACAGCCCATGGATATCCCTGACAGATTAATTAATTTGTTGACCACTATTTATGACCGC 2100
Db 3029 AATGACAGCCCATGGATATCCCTGACAGATTAATTAATTTGTTGACCACTATTTATGACCGC 3088
Qy 2101 CTGAGCAAGAGACAAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAAC 2160
Db 3089 CTGAGCAAGAGACAAACAAATTTGGTCAACGTCCTCTCTGCGTGGATATGTCTGAAC 3148
Qy 2161 TGGCTGCTGAATGTTATGATACGGACGAAACAGGAGGATCCGTCTCTGCTCTTTTAA 2220
Db 3149 TGGCTGCTGAATGTTATGATACGGACGAAACAGGAGGATCCGTCTCTGCTCTTTTAA 3208
Qy 2221 ACTGGCATATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACATACCTTTTCAAG 2280
Db 3209 ACTGGCATATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACATACCTTTTCAAG 3268
2y 2281 CAAGTGGCAAGTTCAACAGATTTTGTACCGACGACGAGCTGGGCTCTCTGCTGATGAT 2340
Db 3269 CAAGTGGCAAGTTCAACAGATTTTGTACCGACGAGCTGGGCTCTCTGCTGATGAT 3328
2y 2341 TCTATCCAAATTCGAAGACAGATTTGGGTGAAGTTGTCATCCCTTTGGGGGCAAGTACATTTGAG 2400
Db 3329 TCTATCCAAATTCGAAGACAGATTTGGGTGAAGTTGTCATCCCTTTGGGGGCAAGTACATTTGAG 3388
2y 2401 C 2401
Db 3389 C 3389

RESULT 2
AR304538
LOCUS AR304538
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 5952)
Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source 1..5952
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN

Query Match 72.3%; Score 1735.6; DB 6; Length 5952;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

2y 628 TGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTAAAGACACTCATAGA 687
Db 2902 TGGAGAAGCATTCATAAAGGGTGTAGAGCGAGAGGCTGCTTTGGAAGAACTCNTAGA 2961
2y 688 TTAAGTCAACAGTTCCTGACCTGGAAAGTTCCTGCTGGCTGTACAGAGCTGAA 747
Db 2962 TTAAGTCAACAGTTCCTGACCTGGAAAGTTCCTGCTGGCTGTACAGAGCTGAA 3021
2y 748 ACAACTGCCAATGCTACAGGATGTCACCGTAAGGAAAGGCTCTAGAAGACTCCAAG 807
Db 3022 ACAACTGCCAATGCTACAGGATGTCACCGTAAGGAAAGGCTCTAGAAGACTCCAAG 3081
Qy 808 GGAGTAAAGAGCTGATGAACAAATGGCAAGCTCCAGGTGAATTCAGCTCACACA 867
Db 3082 GGAGTAAAGAGCTGATGAACAAATGGCAAGCTCCAGGTGAATTCAGCTCACACA 3141

Qy 868 GATGTTATCAACACCTGGATGAAACAGCAAAATCTGAGATCCCTGGAGGTTCC 927
Db 3142 GATGTTATCAACACCTGGATGAAACAGCAAAATCTGAGATCCCTGGAGGTTCC 3201
Qy 928 GATGATCAGTCCCTGTTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTT 987
Db 3202 GATGATCAGTCCCTGTTTACAAAGACGTTTGGATAACATGAACCTCAAGTGGAGTGAACCTT 3261
Qy 988 CGGAAAAAGTCTCTCAACATTTAGGTCCCATTTGGAAGCCAGTTTCTGACCAAGTGAACCGT 1047
Db 3262 CGGAAAAAGTCTCTCAACATTTAGGTCCCATTTGGAAGCCAGTTTCTGACCAAGTGAACCGT 3321
Qy 1048 CTGACACCTTTCTGTCAGGAACCTTCTGCTGTGCTACAGCTGAAAGATGATGAATTAAGC 1107
Db 3322 CTGACACCTTTCTGTCAGGAACCTTCTGCTGTGCTACAGCTGAAAGATGATGAATTAAGC 3381
Qy 1108 CGGAGGACACCTATTGAGAGCGACTTTCCAGACAGTTTCCAGAGCAGAACGATGTACATAGG 1167
Db 3382 CGGAGGACACCTATTGAGAGCGACTTTCCAGAGCAGTTTCCAGAGCAGAACGATGTACATAGG 3441
Qy 1168 GCCTTCAAGAGGGATTTGAAAACTTAAGAACTTGTATCATGATGATCTCTTGAGACTGTA 1227
Db 3442 GCCTTCAAGAGGGATTTGAAAACTTAAAGAACTTGTATCATGATGATCTCTTGAGACTGTA 3501
Qy 1228 CGAATATTTCTGACAGAGCAGCTTTTGGAAAGGACTAGAGAACTCTTACAGAGGCCGAGA 1287
Db 3502 CGAATATTTCTGACAGAGCAGCTTTTGGAAAGGACTAGAGAACTCTTACAGAGGCCGAGA 3561
Qy 1288 GAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTGGGTTTCTAGAAAGCAGGCTGAG 1347
Db 3562 GAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTGGGTTTCTAGAAAGCAGGCTGAG 3621
Qy 1348 GAGCTCAATCTGAGTGGGAAATTTGAACCTCAGCTCCGCTGACCTGGCAGAGAAAATA 1407
Db 3622 GAGTCAATCTGAGTGGGAAATTTGAACCTCAGCTCCGCTGACCTGGCAGAGAAAATA 3681
Qy 1408 GATGAGACCTTTGAAAGACTTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 1467
Db 3682 GATGAGACCTTTGAAAGACTTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGGACCTCAAG 3741
Qy 1468 CTGGCCCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCGCTGGGCGATCTCTCATTTGAC 1527
Db 3742 CTGGCCCAAGCTGAGGTGATCAAGGATCTCTGGCAGCCGCTGGGCGATCTCTCATTTGAC 3801
Qy 1528 TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAA 1587
Db 3802 TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAA 3861
Qy 1588 GAGAACGTGAGCCACGTCATGACCTTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTC 1647
Db 3862 GAGAACGTGAGCCACGTCATGACCTTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTC 3921
Qy 1648 TCACCGTATAACCTCAGCACTCTGGAAGACCTCAACACAGATGGAAGCTTTGCAAGGTG 1707
Db 3922 TCACCGTATAACCTCAGCACTCTGGAAGACCTCAACACAGATGGAAGCTTTGCAAGGTG 3981
Qy 1708 GCCCTCGAGGACCGAGTCAAGGAGCTGTCATGAAGCCCAAGGAGCTTTGGTCCAGCATCT 1767
Db 3982 GCCCTCGAGGACCGAGTCAAGGAGCTGTCATGAAGCCCAAGGAGCTTTGGTCCAGCATCT 4041
Qy 1768 CAGCACTTTCTTTCCAGCTCTGTCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAA 1827
Db 4042 CAGCACTTTCTTTCCAGCTCTGTCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAA 4101
Qy 1828 GTGCCCTACTATATCAACACAGAGACTCAAAACAACTTGTGGGACCATCTCCAAAATGACA 1887
Db 4102 GTGCCCTACTATATCAACACAGAGACTCAAAACAACTTGTGGGACCATCTCCAAAATGACA 4161
Qy 1888 GAGCTCTACAGTCTTTAGCTGACCTGAAATTAATGTGAGATTTCTCAGCTTATAGGACTGCC 1947
Db 4162 GAGCTCTACAGTCTTTAGCTGACCTGAAATTAATGTGAGATTTCTCAGCTTATAGGACTGCC 4221
Qy 1948 ATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGCA 2007

```
Db 4222 ATGAAATCCGAAGATGCGAAGGCCCTTTGCTTGATCTCTTGAGCCTGTGAGCTGCA 4281
Qy 2008 TGTGATGCTTGGACCAAGACACACCTCAAGCAAAATGACAGCCCATGATATCTGCG 2067
Db 4282 TGTGATGCTTGGACCAAGACACACCTCAAGCAAAATGACAGCCCATGATATCTGCG 4341
Qy 2068 ATTATTAAATTTGTGACCACTATTTATGACCGCTGAGCAAGACACAACTTTGGTC 2127
Db 4342 ATTATTAAATTTGTGACCACTATTTATGACCGCTGAGCAAGACACAACTTTGGTC 4401
Qy 2128 AAGCTGCTCTCTGCTGATATGCTGTAAGTGGCTGCTGTAATGTTTATGATACGGGA 2187
Db 4402 AAGCTGCTCTCTGCTGATATGCTGTAAGTGGCTGCTGTAATGTTTATGATACGGGA 4461
Qy 2188 CGAACAGGAGGATCCGTGCTGCTTTTAAACCTGGCATATTTCCCTGCTGTAAGCA 2247
Db 4462 CGAACAGGAGGATCCGTGCTGCTTTTAAACCTGGCATATTTCCCTGCTGTAAGCA 4521
Qy 2248 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGT 2307
Db 4522 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGT 4581
Qy 2308 GACACAGGAGGATCCGTGCTGCTTTTAAACCTGGCATATTTCCCTGCTGTAAGCA 2367
Db 4582 GACACAGGAGGATCCGTGCTGCTTTTAAACCTGGCATATTTCCCTGCTGTAAGCA 4641
Qy 2368 GAAGTTGCTATCTTTGGGGGAGTAACTTGGAGC 2401
Db 4642 GAAGTTGCTATCTTTGGGGGAGTAACTTGGAGC 4675

RESULT 3
AX114289
LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent WO0129243.
ACCESSION AX114289
VERSION AX114289.1 GI:14031259
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Paul,X.L. and Xiao,X.
TITLE Method and vector for producing and transferring trans -spliced peptides
JOURNAL Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION (US)
FEATURES
source 1..5952
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
misc_feature 2897..2898
/notes="S4 junction site"
misc_feature 3198..3199
/notes="S2 junction site"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 72.3% Score 1735.6; DB 6; Length 5952;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 628 TGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTCAAGACACTATAGA 687
Db 2902 TGGAGAGCATTCATAAAGGGTGTGAGCGAGAGGCTGCTTTGGAAGAACTCATAGA 2961
Qy 688 TTACTGCAACAGTTCCCTCGGACCTGGAAGATTTCTTGGCTGGCTTACAGAGCTGAA 747
Db 2962 TTACTGCAACAGTTCCCTCGGACCTGGAAGATTTCTTGGCTGGCTTACAGAGCTGAA 3021
```

```
Qy 748 ACACCTGCCAATGTCTTACAGGATGCTACCCGTAGAAAGGCTCTCTAGAGACTCCAA 807
Db 3022 ACACCTGCCAATGTCTTACAGGATGCTACCCGTAGAAAGGCTCTCTAGAGACTCCAA 3081
Qy 808 GGAGTAAAGAGCTGTATGAAACCAATGCAAGACCTTCAAGGTGAAATTTGAAGCTCAC 867
Db 3082 GGAGTAAAGAGCTGTATGAAACCAATGCAAGACCTTCAAGGTGAAATTTGAAGCTCAC 3141
Qy 868 GATGTTTATCACAACTGAGTGAACCAAGACCAAAATCTCTGAGATCCCTTGGAGGTTCC 927
Db 3142 GATGTTTATCACAACTGAGTGAACCAAGACCAAAATCTCTGAGATCCCTTGGAGGTTCC 3201
Qy 928 GATGATCAGTCTCTGTATCAAGACCTTTGGATTAACATGAATTTCAAGTGGAGTGA 987
Db 3202 GATGATCAGTCTCTGTATCAAGACCTTTGGATTAACATGAATTTCAAGTGGAGTGA 3261
Qy 988 CGGAAAGTCTCTCAACATTTAGTGTCCATTTGGAAGCCAGTTCTGACAGTGGAGCGT 1047
Db 3262 CGGAAAGTCTCTCAACATTTAGTGTCCATTTGGAAGCCAGTTCTGACAGTGGAGCGT 3321
Qy 1048 CTGCACCTTTCTCTGCAAGGAACTTTCTGGTGTGCTACAGCTGAAAGATGATGAATTA 1107
Db 3322 CTGCACCTTTCTCTGCAAGGAACTTTCTGGTGTGCTACAGCTGAAAGATGATGAATTA 3381
Qy 1108 CGGAGGCACTATTTGGAGCGACTTTTCCAGAGTTTCCAGAGCAGACGATGTACATAG 1167
Db 3382 CGGAGGCACTATTTGGAGCGACTTTTCCAGAGTTTCCAGAGCAGACGATGTACATAG 3441
Qy 1168 GCCTTCAAGAGGAAATTTGAAACTTAAAGACCTGTATCATGAGTACTCTTTGAGACTGA 1227
Db 3442 GCCTTCAAGAGGAAATTTGAAACTTAAAGACCTGTATCATGAGTACTCTTTGAGACTGA 3501
Qy 1228 CGAATATTTCTGACAGAGCAGCTTTTGGAGGACTAGAGAACTCTACAGAGCCACAGA 1287
Db 3502 CGAATATTTCTGACAGAGCAGCTTTTGGAGGACTAGAGAACTCTACAGAGCCACAGA 3561
Qy 1288 GAGCTGCTCTCTGAGGAGAGCCGAGATGTCACTCGGCTTCTACGAAAGAGGCTGAG 1347
Db 3562 GAGCTGCTCTCTGAGGAGAGCCGAGATGTCACTCGGCTTCTACGAAAGAGGCTGAG 3621
Qy 1348 GAGCTCAATATCTGAGTGGGAAATTTGAACCTGCACCTCGCTGACTGGCAGAGAAATA 1407
Db 3622 GAGCTCAATATCTGAGTGGGAAATTTGAACCTGCACCTCGCTGACTGGCAGAGAAATA 3681
Qy 1408 GATGAGACCTTTGAGAGACTCCAGAGACTTCAAGAGCCACCGATGAGCTGGACCTCA 1467
Db 3682 GATGAGACCTTTGAGAGACTCCAGAGACTTCAAGAGCCACCGATGAGCTGGACCTCA 3741
Qy 1468 CTGGCCCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCGTGGCGATCTCTCATTTGAC 1527
Db 3742 CTGGCCCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCGTGGCGATCTCTCATTTGAC 3801
Qy 1528 TCTCTCAAGATCACTCTGAGAAAGTCAAGGACTTTCGAGGAGAAATTTGGCCTCTGAAA 1587
Db 3802 TCTCTCAAGATCACTCTGAGAAAGTCAAGGACTTTCGAGGAGAAATTTGGCCTCTGAAA 3861
Qy 1588 GAGAACTGAGGACCACTCAATGACTTCTCGCCAGCTTACCACTTTGGGCATTCAGCTC 1647
Db 3862 GAGAACTGAGGACCACTCAATGACTTCTCGCCAGCTTACCACTTTGGGCATTCAGCTC 3921
Qy 1648 TCACCGTATAACCTCAGCACTCTGGAAGACTTGAACCAACAGATGGAAGCTTCTGAGGTG 1707
Db 3922 TCACCGTATAACCTCAGCACTCTGGAAGACTTGAACCAACAGATGGAAGCTTCTGAGGTG 3981
Qy 1708 GCCGTGAGGACCGAGTCAAGGCTGATGAGAGCCACAGGAGCTTTGGTCCAGGATCT 1767
Db 3982 GCCGTGAGGACCGAGTCAAGGCTGATGAGAGCCACAGGAGCTTTGGTCCAGGATCT 4041
Qy 1768 CAGCACTTTCTTTTCCAGCTGTCTCCAGGCTTCCCTGGGAGAGAGCCATCTCGCCAAACAA 1827
Db 4042 CAGCACTTTCTTTTCCAGCTGTCTCCAGGCTTCCCTGGGAGAGAGAGCCATCTCGCCAAACAA 4101
```

QY 1828 GTGCCCTACTATCAACACGAGACTCAAAACATTTGCTGGGACCATCCAAAATGACA 1887
DB 4102 GTGCCCTACTATATCAACACGAGACTCAAAACATTTGCTGGGACCATCCAAAATGACA 4161
QY 1888 GAGCTCTACAGTCTTTAGCTGAGCTGAAATATGTGAGATCTCAGCTTATAGAGTGGC 1947
DB 4162 GAGCTCTACAGTCTTTAGCTGAGCTGAAATATGTGAGATCTCAGCTTATAGAGTGGC 4221
QY 1948 ATGAACCTCCGAAGACTGCAAGAGGCGCTTTGCTTGGATCTCTGAGCCCTGTGAGCTGCA 2007
DB 4222 ATGAACCTCCGAAGACTGCAAGAGGCGCTTTGCTTGGATCTCTGAGCCCTGTGAGCTGCA 4281
QY 2008 TGTGATGCTTGGACGACGACCAACCTCAAGCAAAATGACGAGCCCATGGATATCTTCGAG 2067
DB 4282 TGTGATGCTTGGACGACGACCAACCTCAAGCAAAATGACGAGCCCATGGATATCTTCGAG 4341
QY 2068 ATTATTAATCTTTGACCACTATTATGACCGCTGAGCAAGAGGCAAGACACAAATTTGGTC 2127
DB 4342 ATTATTAATCTTTGACCACTATTATGACCGCTGAGCAAGAGGCAAGACACAAATTTGGTC 4401
QY 2128 AACGTCCCTCTCTCGTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGA 2187
DB 4402 AACGTCCCTCTCTCGTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGA 4461
QY 2188 CGACAGGAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
DB 4462 CGACAGGAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 4521
QY 2248 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGT 2307
DB 4522 CATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGT 4581
QY 2308 GACAGCGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCACAGACAGTTGGGT 2367
DB 4582 GACAGCGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCACAGACAGTTGGGT 4641
QY 2368 GAAGTTGATCTTTGGGGGAGTAAACATTGAGC 2401
DB 4642 GAAGTTGATCTTTGGGGGAGTAAACATTGAGC 4675

RESULT 4
AX538622
LOCUS AX538622 8689 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..8689
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"

BASE COUNT 2721 a 1804 c 1861 g 2303 t
ORIGIN

Query Match 72.3%; Score 1735.6; DB 6; Length 8689;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 628 TGGCAACACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAGACATCTATAGA 687
DB 2948 TGGAGAGCAATTCATAAAGGGTGAGTGGAGAGGCTCTTTGGAGAACTCATAGA 3007

QY 688 TTACTGCAACAGTTCCTCCCTGAGCCTGCAAAAGTTTCTTGTGCTGGCTTACAGAACTGAA 747
DB 3008 TTACTGCAACAGTTCCTCCCTGAGCCTGCAAAAGTTTCTTGTGCTGGCTTACAGAACTGAA 3067
QY 748 ACACTGCAACATGTCTCTACAGATGCTACCCGTAAAGAAAGCTCCTTAGAGACTCCAAAG 807
DB 3068 ACACTGCAACATGTCTCTACAGATGCTACCCGTAAAGAAAGCTCCTTAGAGACTCCAAAG 3127
QY 808 GAGTAAAGAGCTCATGAAACAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACA 867
DB 3128 GAGTAAAGAGCTCATGAAACAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACA 3187
QY 868 GATGTTTATCAACCTGGATGAAACAGCCAAATAATCTTGAGATCCCTGGAAGTTCC 927
DB 3188 GATGTTTATCAACCTGGATGAAACAGCCAAATAATCTTGAGATCCCTGGAAGTTCC 3247
QY 928 GATGATGAGTCTCTGTTACAAAGAGCTTTGGATGACATGAACTCAAGTGAGTCAACTT 987
DB 3248 GATGATGAGTCTCTGTTACAAAGAGCTTTGGATGACATGAACTCAAGTGAGTCAACTT 3307
QY 988 CGGAAAAAGTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTCTTGACCAGTGAAGCGT 1047
DB 3308 CGGAAAAAGTCTCTCAACATTAGGTCCCAATTTGGAAGCCAGTTCTTGACCAGTGAAGCGT 3367
QY 1048 CTGCACTTTCTCTGACAGAACTTCTGCTGAGCTAGAGCTGAAAGATGATGAATTAAGC 1107
DB 3368 CTGCACTTTCTCTGACAGAACTTCTGCTGAGCTAGAGCTGAAAGATGATGAATTAAGC 3427
QY 1108 CGGCAAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGG 1167
DB 3428 CGGCAAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGG 3487
QY 1168 GCCTTCAAGAGGAATTCGAAACCTTAAAGAACTGTATCATGAGTACTCTTTGAGACTGTA 1227
DB 3488 GCCTTCAAGAGGAATTCGAAACCTTAAAGAACTGTATCATGAGTACTCTTTGAGACTGTA 3547
QY 1228 CGAATATTTCTGACAGACAGCCTTTTGAAGGACTTAGAGAACTTCTACAGAGGCCAGA 1287
DB 3548 CGAATATTTCTGACAGACAGCCTTTTGAAGGACTTAGAGAACTTCTACAGAGGCCAGA 3607
QY 1288 GAGCTGCCCTCTGAGGAGAGAGCCAGAAATCTCACTCGGCTTCTACGAAAGCAGCTGAG 1347
DB 3608 GAGCTGCCCTCTGAGGAGAGAGCCAGAAATCTCACTCGGCTTCTACGAAAGCAGCTGAG 3667
QY 1348 GAGTCAATGCTGAGTGGGAAATTTGAACTGCTGACTCCGCTGACTGGCAGAGAAATA 1407
DB 3668 GAGTCAATGCTGAGTGGGAAATTTGAACTGCTGACTCCGCTGACTGGCAGAGAAATA 3727
QY 1408 GATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAG 1467
DB 3728 GATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAG 3787
QY 1468 CTGCGCCAAAGCTGAGGTGATCAAGGATCTCTGGGAGCCCGTGGGCGATCTCTCATTTGAC 1527
DB 3788 CTGCGCCAAAGCTGAGGTGATCAAGGATCTCTGGGAGCCCGTGGGCGATCTCTCATTTGAC 3847
QY 1528 TCTCTCCAAGATCACCTCGAGAAAGTCAAGSCATTCGAGAGAAATTTGGCCCTCTGAAA 1587
DB 3848 TCTCTCCAAGATCACCTCGAGAAAGTCAAGSCATTCGAGAGAAATTTGGCCCTCTGAAA 3907
QY 1588 GAGAACGTGAGCCACGTCAATGACCTTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTC 1647
DB 3908 GAGAACGTGAGCCACGTCAATGACCTTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTC 3967
QY 1648 TCACCGTATACCTCAGCACTCTGGAAGCTGAAACACCCAGATGGAAGCTTCTCGAGTG 1707
DB 3968 TCACCGTATACCTCAGCACTCTGGAAGCTGAAACACCCAGATGGAAGCTTCTCGAGTG 4027
QY 1708 GCGGTCCAGGACCGAGTCAAGCTGAGTGAAGCCCAAGGAGCTTTGGTCCAGCATCT 1767
DB 4028 GCGGTCCAGGACCGAGTCAAGCTGAGTGAAGCCCAAGGAGCTTTGGTCCAGCATCT 4087
QY 1768 CAGCACTTTCTTCCAGCTCTGTCCAGGCTCCCTTGGGAGAGAGCCATCTCGCCAAACAA 1827


```
Db 6782 GCCGTGAGGACCGAGTTCAGCGACTGCATGAAAGCCACAGGAGACTTTGGTCAGCATCT 6841
Qy 1768 CAGCACTTTCTTTCCAGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1827
Db 6842 CAGCACTTTCTTTCCAGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 6901
Qy 1828 GTGCGCTACTATATCAACACGAGAGACTCAACAACTTGTCTGGAGACCATCCAAATGACA 1887
Db 6902 GTGCGCTACTATATCAACACGAGAGACTCAACAACTTGTCTGGAGACCATCCAAATGACA 6961
Qy 1888 GAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTCTCAGATTCTCAGCTTATAGCACTGCC 1947
Db 6962 GAGCTCTACCAAGTCTTTAGCTGACCTGAATAATGTCTCAGATTCTCAGCTTATAGCACTGCC 7021
Qy 1948 ATGAACTCCGAGACTGCGAAGAGCCCTTTGTCTGGATCTCTTGAGCCCTGTGAGCTGCA 2007
Db 7022 ATGAACTCCGAGACTGCGAAGAGCCCTTTGTCTGGATCTCTTGAGCCCTGTGAGCTGCA 7081
Qy 2008 TGTGATGCCCTTGGACACGACCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCGAG 2067
Db 7082 TGTGATGCCCTTGGACACGACCAACCTCAAGCAAAATGACCCAGCCCATGGATATCCTGCGAG 7141
Qy 2068 ATTATTAATTTGTTGACCACTATTATGACCGCTGAGCAAGACAGACCAACAATTTGGTC 2127
Db 7142 ATTATTAATTTGTTGACCACTATTATGACCGCTGAGCAAGACAGACCAACAATTTGGTC 7201
Qy 2128 AACGTCCTCTCGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGA 2187
Db 7202 AACGTCCTCTCGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGA 7261
Qy 2188 CGAACAGGAGGATCCGTGTCCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
Db 7262 CGAACAGGAGGATCCGTGTCCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 7321
Qy 2248 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGT 2307
Db 7322 CATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGT 7381
Qy 2308 GACCAGCGCAGGCTGGGCTCTCTGCAATGTTTATCCAAATTCACAGAGTTGGCT 2367
Db 7382 GACCAGCGCAGGCTGGGCTCTCTGCAATGTTTATCCAAATTCACAGAGTTGGCT 7441
Qy 2368 GAAGTTGCATCTTTGGGGCAGTAACATTGAGC 2401
Db 7442 GAAGTTGCATCTTTGGGGCAGTAACATTGAGC 7475

RESULT 6
AX538627 12057 bp DNA linear PAT 23-NOV-2002
LOCUS Sequence 47 from Patent WO0229056.
DEFINITION AX538627
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
    Location/Qualifiers
        1..12057
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
            /note="Synthetic"
BASE COUNT 4020 a 2456 c 2790 g 2791 t
ORIGIN
Query Match 72.3%; Score 1735.6; DB 6; Length 12057;
Best Local Similarity 98.6%; Pred. No. 0;
```

```
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 628 TGGGCAAAACATCTGTAGATGACAGAAAGACCGCTGGGTCTCTTTTACAAGACACTCATAGA 687
Db 8216 TGGGAAGCAATTCATAAAAGGGTGAGTCAGCGAGAGGCTGCTTTTGGGAAGAACTCATAGA 8275
Qy 688 TTACTGCAACAGTTCCCTCGGACCTGGAAAGTTTCTTGCTCGCTGTACAGAGCTGAA 747
Db 8276 TTACTGCAACAGTTCCCTCGGACCTGGAAAGTTTCTTGCTCGCTGTACAGAGCTGAA 8335
Qy 748 ACAACTGCAATGCTTACAGAGTCTACCGTAAGGAAGGCTCCTAGAAGACTCCAAAG 807
Db 8336 ACAACTGCAATGCTTACAGAGTCTACCGTAAGGAAGGCTCCTAGAAGACTCCAAAG 8395
Qy 808 GAGTAAGAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 867
Db 8396 GAGTAAGAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 8455
Qy 868 GATGTTTATCAACACCTGGATGAAACAGCAAGCAAAATCCTGAGATCCCTGGAAGTTCC 927
Db 8456 GATGTTTATCAACACCTGGATGAAACAGCAAGCAAAATCCTGAGATCCCTGGAAGTTCC 8515
Qy 928 GATGATGCACTCTGTTTACAAAGAGCGTTTGGATTAACATGAACCTTCAAGTGGAGTGAAT 987
Db 8516 GATGATGCACTCTGTTTACAAAGAGCGTTTGGATTAACATGAACCTTCAAGTGGAGTGAAT 8575
Qy 988 CGGAAAAGTCTCTCAACATTTAGGTCCCATTTTGAAGCCAGTTCTGACCAGTGGAAAGCT 1047
Db 8576 CGGAAAAGTCTCTCAACATTTAGGTCCCATTTTGAAGCCAGTTCTGACCAGTGGAAAGCT 8635
Qy 1048 CTGCACCTTTCTCTGCAAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
Db 8636 CTGCACCTTTCTCTGCAAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8695
Qy 1108 CGGAGGCACTTATGAGGCGCACTTTCCAGCACTTCAGAAAGCAGAACGATGTACATAGG 1167
Db 8696 CGGAGGCACTTATGAGGCGCACTTTCCAGCACTTCAGAAAGCAGAACGATGTACATAGG 8755
Qy 1168 GCCTTCAAGGGGAATTCGAAACCTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 1227
Db 8756 GCCTTCAAGGGGAATTCGAAACCTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 8815
Qy 1228 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGCACTAGAGAACTCTACAGAGGCCAGA 1287
Db 8816 CGAATATTTCTGACAGAGCAGCCTTTTGGAAAGCACTAGAGAACTCTACAGAGGCCAGA 8875
Qy 1288 GAGCTGCTCTCTGAGGAGAGAGCCAGCAATGTCACTCGGCTTCTACGAAAGCGGCTGAG 1347
Db 8876 GAGCTGCTCTCTGAGGAGAGAGCCAGCAATGTCACTCGGCTTCTACGAAAGCGGCTGAG 8935
Qy 1348 GAGGTCAATACTGAGTGGGAAATAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA 1407
Db 8936 GAGGTCAATACTGAGTGGGAAATAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATA 8995
Qy 1408 GATGAGACCTTTGAAAGACTCCAGGAACTTCAGAGAGCCACGAGTGAAGTGGACCTCAAG 1467
Db 8996 GATGAGACCTTTGAAAGACTCCAGGAACTTCAGAGAGCCACGAGTGAAGTGGACCTCAAG 9055
Qy 1468 CTGCGCCAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGATCTCTCATTTGAC 1527
Db 9056 CTGCGCCAGCTGAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGATCTCTCATTTGAC 9115
Qy 1528 TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAA 1587
Db 9116 TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAA 9175
Qy 1588 GAGAACGTGAGCCACCTCATGACCTTGTGCGCAGCTTACCACTTTGGCATTTCAGCTC 1647
Db 9176 GAGAACGTGAGCCACCTCATGACCTTGTGCGCAGCTTACCACTTTGGCATTTCAGCTC 9235
Qy 1648 TCACCGTATACCTTCAGCACTCTGGAAGAGCTGAACACACAGATGGAAGCTTCTGCAGGTG 1707
Db 9236 TCACCGTATACCTTCAGCACTCTGGAAGAGCTGAACACACAGATGGAAGCTTCTGCAGGTG 9295
```

QY	1708	GCCTGCGAGGACCGAGTCTCAGGAGCTGCATGAAGCCACAGGAGCATTTGGTCCAGCATCT	1767
Db	9296	GCCTGCGAGGACCGAGTCTCAGGAGCTGCATGAAGCCACAGGAGCATTTGGTCCAGCATCT	9355
QY	1768	CAGCAGCTTTCTTCCAGCTCTCTCAGGCTCTCTCAGGCTCTCTGGAGAGAGCCATCTCGCCAAACAAA	1827
Db	9356	CAGCAGCTTTCTTCCAGCTCTCTCAGGCTCTCTCAGGCTCTCTGGAGAGAGCCATCTCGCCAAACAAA	9415
QY	1828	GTGCCCTACTATATCAACACAGAGACTCAACAACTTGTGGGACCATCCCAAAATGACA	1887
Db	9416	GTGCCCTACTATATCAACACAGAGACTCAACAACTTGTGGGACCATCCCAAAATGACA	9475
QY	1888	GAGCTCTACAGCTTTTACTGACCTGTAATATGTCAGATTTCTAGCTATAGGACTGCC	1947
Db	9476	GAGCTCTACAGCTTTTACTGACCTGTAATATGTCAGATTTCTAGCTATAGGACTGCC	9535
QY	1948	ATGAAGAACTCCGAGAGCTGCAGAGGCGCTTCTCTGGATCTCTTGCAGCTGTGCAGTGCA	2007
Db	9536	ATGAAGAACTCCGAGAGCTGCAGAGGCGCTTCTCTGGATCTCTTGCAGCTGTGCAGTGCA	9595
QY	2008	TGTGATGCTCTTGACACGACCAACCTCAAGCAAAATGACACCGCCATCGATATCTCGAG	2067
Db	9596	TGTGATGCTCTTGACACGACCAACCTCAAGCAAAATGACACCGCCATCGATATCTCGAG	9655
QY	2068	ATTATTAAATGTTTGACCACTATTATGACCGCTCTGGAGAGAGACACAAATTTGGTC	2127
Db	9656	ATTATTAAATGTTTGACCACTATTATGACCGCTCTGGAGAGAGACACAAATTTGGTC	9715
QY	2128	AACGTCCTCTCTGCTGATATGCTCTGAACTGGCTGCTGAATGTTTATGATACGGGA	2187
Db	9716	AACGTCCTCTCTGCTGATATGCTCTGAACTGGCTGCTGAATGTTTATGATACGGGA	9775
QY	2188	CGAAGGAGAGATCGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA	2247
Db	9776	CGAAGGAGAGATCGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA	9835
QY	2248	CATTTCGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGT	2307
Db	9836	CATTTCGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGT	9895
QY	2308	GACGAGGAGGCTGGGCTCTCTTCTGATGATTTATCCAAATTCGAAGACAGTTGGGT	2367
Db	9896	GACGAGGAGGCTGGGCTCTCTTCTGATGATTTATCCAAATTCGAAGACAGTTGGGT	9955
QY	2368	GAAGTTGCATCTCTTGGGGCGAGTACATTTGAGC	2401
Db	9956	GAAGTTGCATCTCTTGGGGCGAGTACATTTGAGC	9989
RESULT 7			
HSDMDR			
LOCUS	HSDMDR	12446 bp	mRNA linear PRI 12-SEP-1993
DEFINITION	Human mRNA for dystrophin.		
ACCESSION	X14298		
VERSION	X14298.1	GI:30845	
KEYWORDS	Dmd gene; Duchenne muscular dystrophy; dystrophin.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	Rosenthal, A., Speer, A., Billwitz, H., Cross, G.S., Forrest, S.M. and Davies, K.E.		
TITLE	Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous		
JOURNAL	Nucleic Acids Res. 17 (13), 5391 (1989)		
MEDLINE	89345106		
PUBMED	266885		
REFERENCE	1 (bases 1 to 12446)		
AUTHORS	Rosenthal, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften		

der DDr, Zentralinstitut fuer Molekularbiologie, Robert-Roessele Str.10, 1115 Berlin Buch, DDR
see also M18533 and M2050 for Dmd segs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.

FEATURES

source

1. .12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X chromosomal, Xp21."
/issue_type="muscle"
/dev_stage="fetal and adult."

99. .11156
/notes="dystrophin (AA 1 - 3685)"

CDS

/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/translation="MLWEEVEDCYEREDVQKTFKWNQAQSKFGKHIEIENLSD
QDRLDLEGLTGQKPKRGSTRVHALNNVKNALRVLQNNVDLVNIGSTDIVD
NKHLLGLTWNILHMQVNMKNVAGLQPTNSBKILLWSVRQTRNTPQVNVNFT
TWSGDLNALHSHRPDLFDWNVVQCSATQRLHAFNARTYQLGIEKLDPEDV
DTTPPKKILMTITSLFQVLSIAIEQVEMLPKPPKVTKEBHPQLHQMYSQ
QITSLAQGVERTSPKPKSVAYTQAAYVTSPTSPSPSPSPSPSPSPSPSPSP
MESEVNLDRYOTALBEVLWMLLSAEDTLQAQCEIENDVEVKDQFHTHEGYMDLT
QQRVNLGLGSKLGTGKLSDEDETEVQEQNNLNSRECLRVASMEKQSNLHVLM
DQONQKLENDLWLTKEERTKMEEPPLGDEDLKQVQKHQLQDLEGEQVRVN
SLTHMVVVDESDDHATALEEQKLGDRWANI CRWTEDRWVLLDILLKQRLTE
ECLFSALISEKEDAVNKIHTTFKQDNEMLSLQKLAJLADLEKQSGMKGLYSIK
QDLSTLNKNSVTQKTEAWLNFARQDNMLVQKLEKSTAQISQAVTTPSPSTTVM
ETVVTTPTEOILVKAQBELPPPKPKRQITVDSERKRLDITELHWTISEA
VLSPEPAI FRKGEFSDLEKVNALEREKFKLODASRSGOALVEQMVNENVA
DIKQASQOLNSRWIEFCOLLSERLNWLEYQNNIIAFYNQLOQLQEQMTTAEWNLKIO
PTTSPPTAKSLQKICXENVRLSQLOPQIRLKIQTALKEKQSGPMFLDADVFAP
THKQVDFDQVQAREKELQITFDLTPPMRYQETMSAIRTVQOQSTKUSIPOLSVTDY
ELMEQVGLQALQSLQSGSLYSLTVTKMSKAPSEISRYQSEFEFEIEGRKQ
KLSQVLEHCKLEQNKIQNHIOITLKKMAEVDVFLKEBLNTQWMDHCOOY
KCRILVSDIOTIPLSNVNEGOKI KNEAPEFPASLEBELKELNTQWMDHCOOY
ARKEALGDEKLEKTSLOKLSLSEHWMQTAEEYLERDFEYKTPDELQKAPREMCRK
EAAQKAKVLLTSVNSVIAQAPVQAQALKELETLITNYQWCTELNGCKTLE
EYVACWHELLSYLEKANKWLNVEFLKTTENIPGAAEIESEVLSLENLHRSDENP
NTRILAQITDGVMDLEINEELTFNSRHELHEEAVRQKLEQSIQSAQETNS
LHLISQIFIDKQLAAVIAADQAPQEAQISDLTSHIEISLEMKHNGKQEA
AORLSQIDVAOKLQDVMKFLPKFQKPNQRLQESKMLDEVMKHLPALETQSE
QEVQSNLHCNLYKSLSEKSEVEMVLTGROIQVOKTENKPELDERVLTALXHY
NELKAVTERKQKLEKLSRKMKNVLTJLAATDMELTKESEVAGWPSNDSE
VAWGKATQKEIEKQVHLKSTIEVGRALKTVLGKKEITVEDKLSLNSNWIATVRAE
EMLNLLLEYKHMTFDQNDVHTIKWIOADTLLDESEKPKQKQEDVLRKRAELND
IRPKVDSTRQANLMAHGDHCRKLVFQISELNHRFAAISHRIKTKGASIPLEKE
QFNSDIQKLEPLEABIQGVNLEEDFNKMNEDNEGTVELLQRLQRLQRIITDER
KSEIIKIQQLLQTHNALKDLRSRRKKAIEISHQYQKQADLLKCLDIDIEKLL
ASLPEPRDERKIEIDREIKKELNARROAGLSEDDGAAVETQIOLSKWRE
LISKPAQFLNPAQIHTVREEMTMEDMELEISYVPSVLTETITVUSQALREVO
LNAADQACAFEDFKQESLKNIDSLQSSGRIDIHSHKTAALQASATPVERVKL
QALSQLDFQWKNVNMQRDRFDRSVKRRRHHYDIKIFNQMLTAEQELKQITQI
PENHEAKYKWLKELODQIGQRTVRLNATGEEIIQQSKTASILAEQKLSGLN
RWQEVCKLSDEKLEQNLSEFFQDLNEFVLEADNIAISIPLEPGEQQLKE
KLEOVKLVLEPLRQGLKQNLNTEGPPVLSAPISPEODKLNKLTQNLQWIKV
LALPEKQIEAQIDKLGLEKLEBOLNHLILSPINQLEIYNQNPQEQPFD
VETEIAVQAQPDVEELISKQHLKKEKPAQPVKSKLELSSSWKAVPALNPAWT
KQDPAAGTITGASPTQITVITVPTVGTATISKLEPSSMLMEVPAALNPAWT
EUTDLSLIDQIKQVNMVDEINEMIIKQATMQDLEQRRQKLEQLEKLEKLEKLE
KNTSQEAETIITDRIERIQWDEVEQLNQRQNLNEMKDSQWLEAKESQVLE
GOARAKLSEWEGPVTDAIOKKTETKQALDKRQWQTNVDVANDLALKLRYDSAD
DTRKVMITENINAGRSVTHKVSEREALETHRLQOFPDLDEKFLAWLFEATTA
NVLODARTKRELLESKGVKELMKQWQDLOQIEIETHVDVYHNLDSOKILSLBESD
DAVLLORLDNNNFKSELKSLNIRSHLEASSQWELHLSLOELVWLKODEL
SQAPIGGDFPAPQKNDVHRAPKELKTKGPIVINSTLETETVRIETEQPLEGLEKLYQ
EPRELPEERAQNVTRLKQAEVNTWEKLNLSADQWQKIDETLERKLEQLQRAID
ELDKLQRAEVIKGSQWQVGDILLIDSLQDLEKVKALKEIAPLEKLVSHVNDLARQ
TTIGIQLSPYNLSTEDLNTRWKLQVAVEDVRQLHEARDLFGPASQFLSTSVQGP
WEAISPNKVPVYINHTQTTCWDHPKMTLEYQSLADLNNVRFSAYRTAMKRLRLQKA

LCGLDLSLSAACDALDHNKLNKQNDQPMDDILOINCLITTYIDRLQBNHNLVNVPLCVD
MCNLNLLNVDTGRTGRIRVLSFKTGLISLCKAHLEDKRYLFKQVASTGFCDDRRLL
GLLHSDIQIPLQGEVASFSGNIIFSVASCQFANKRFEIEAALFLDMWRLEPQSM
VWLPVLRHVAATTAHQDQKNCIKCKPIIGFRYSLKHFNVDICOSCFPSVRKAGH
KMHVPMEYCTPTTSGDVRDFAKLVKNKPRTRYFAKHPRMGYLPVQTVLGRDNMET
PVTLLNFWPVDSPASSPOLSHDDTHSRLEHYASRLAEMENSGSLVNDISIPNSBID
DEHLLHQYQCSLNQDPSPLSQSPSPAQIILSLESEBGERLELADLEENRNLOQAY
DRLKQHEHGLSPLSPPEMPTSPQSPRDAELIAEAKLLROHKHLEARNMOILEDH
NKQESOLHLRLQLLPQABAKVNGTIVTSSPSTSLORSDSQPMPLRVVGSQTSDSM
GBEDLLSPQDSTGHEVNEQNLNPFPSRGRNTPGKPMREDTM

BASE COUNT 4135 a 2524 c 2876 g 2911 t

ORIGIN

Query Match 72.38; Score 1735.6; DB 9; Length 12446;
Best Local Similarity 98.64; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy	628	TGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACATCATAGA	687
Db	8106	TGGAGAGCATTCATAAAGGGTGAGTGAGGAGAGGCTGCTTTGGAAGNAATCATAGA	8165
Qy	688	TTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAGCTGAA	747
Db	8166	TTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAGCTGAA	8225
Qy	748	ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGCTCTAGNAGACTCCAG	807
Db	8226	ACAACTGCCAATGTCCTACAGGATGCTACCCGTAAGGAAGGCTCTAGNAGACTCCAG	8285
Qy	808	GGAGTAAAAGAGCTGATGAAACAAATGGAAGACCTCCAGGTGAAATTTGAAGCTCACACA	867
Db	8286	GGAGTAAAAGAGCTGATGAAACAAATGGAAGACCTCCAGGTGAAATTTGAAGCTCACACA	8345
Qy	868	GATGTTTATCACACCTGGATGAACAGACCAAAATCCTGAGATCCCTGGNAGTTCC	927
Db	8346	GATGTTTATCACACCTGGATGAACAGACCAAAATCCTGAGATCCCTGGNAGTTCC	8405
Qy	928	GATGATGAGTCTGTTTACAAAGACGTTTGATGATAACATGAATTTCAAGTGGAGTGAAT	987
Db	8406	GATGATGAGTCTGTTTACAAAGACGTTTGATGATAACATGAATTTCAAGTGGAGTGAAT	8465
Qy	988	CGHAAAAGTCTCTCAACATAGTCCATTTGGAGCCAGTTCTGACCAAGTGGAAAGCT	1047
Db	8466	CGHAAAAGTCTCTCAACATAGTCCATTTGGAGCCAGTTCTGACCAAGTGGAAAGCT	8525
Qy	1048	CTGCACCTTTCTCTGACAGGAATCTGCTGTGCTACAGCTGAAGATGATGAATTAAGC	1107
Db	8526	CTGCACCTTTCTCTGACAGGAATCTGCTGTGCTACAGCTGAAGATGATGAATTAAGC	8585
Qy	1108	CGGACGACCTATTGAGGCGACTTTCCAGCAGTTTCAGAACGAGACGATGTACATAGG	1167
Db	8586	CGGACGACCTATTGAGGCGACTTTCCAGCAGTTTCAGAACGAGACGATGTACATAGG	8645
Qy	1168	GCCTTCAAGAGGGAATGAAATCTAAGACCTGTAATCATGAGTACTCTTGAGACTGTA	1227
Db	8646	GCCTTCAAGAGGGAATGAAATCTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTA	8705
Qy	1228	CGAATATTTCTGACAGACGACCTTTTGGAGGACTAGAGAACTCTPACAGGAGGCCAGA	1287
Db	8706	CGAATATTTCTGACAGACGACCTTTTGGAGGACTAGAGAACTCTPACAGGAGGCCAGA	8765
Qy	1288	GAGCTGCTCTCTGAGGAGAGGCGAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAG	1347
Db	8766	GAGCTGCTCTCTGAGGAGAGGCGAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAG	8825
Qy	1348	GAGTCAATACTAGTGGGAAAATTTGAACCTGCATCTCCGCTGACTGGCAGAGAAAATA	1407
Db	8826	GAGTCAATACTAGTGGGAAAATTTGAACCTGCATCTCCGCTGACTGGCAGAGAAAATA	8885
Qy	1408	GATGAGACCTTGAAGACTTCCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAG	1467
Db	8886	GATGAGACCTTGAAGACTTCCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAG	8945

Qy	1468	CTGCGCCAAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGATCTCTCTCATTTGAC	1527
Db	8946	CTGCGCCAAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGATCTCTCTCATTTGAC	9005
Qy	1528	TCTCTCCAAGATCACTCTGAGAAAGTCAAGGCACTTCGAGAGGAAATTTGGCCTCTGAAA	1587
Db	9006	TCTCTCCAAGATCACTCTGAGAAAGTCAAGGCACTTCGAGAGGAAATTTGGCCTCTCTGAAA	9065
Qy	1588	GAGAACTGTGAGCCACGCTCAATGACTTCTGCTCGCAGCTTACCCTTTTGGGCAATTCAGCTC	1647
Db	9066	GAGAACTGTGAGCCACGCTCAATGACTTCTGCTCGCAGCTTACCCTTTTGGGCAATTCAGCTC	9125
Qy	1648	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTCGAGGTG	1707
Db	9126	TCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCCAGATGGAAGCTTCTCGAGGTG	9185
Qy	1708	CCCGTCGAGGACCGAGTCAGGACGCTGATGAAGCCCAACAGGACCTTTGGTCCAGCATCT	1767
Db	9186	CCCGTCGAGGACCGAGTCAGGACGCTGATGAAGCCCAACAGGACCTTTGGTCCAGCATCT	9245
Qy	1768	CAGCACTTTCTTTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAA	1827
Db	9246	CAGCACTTTCTTTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAA	9305
Qy	1828	GTGCGCTACTATATCAACCCAGACACTCAACAACTTGTGGGACCATCCCAAAATGACA	1887
Db	9306	GTGCGCTACTATATCAACCCAGACACTCAACAACTTGTGGGACCATCCCAAAATGACA	9365
Qy	1888	GAGCTTACCAAGTCTTTAGTGACCTGAATAATGTCAGATTTCTCAGCTTATAGGACTGCC	1947
Db	9366	GAGCTTACCAAGTCTTTAGTGACCTGAATAATGTCAGATTTCTCAGCTTATAGGACTGCC	9425
Qy	1948	ATGAAACTCCGAAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCA	2007
Db	9426	ATGAAACTCCGAAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCA	9485
Qy	2008	TGTGATGCTTTGGACCCAGCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG	2067
Db	9486	TGTGATGCTTTGGACCCAGCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG	9545
Qy	2068	ATTATTAATTTGTTTGAAGCACTATTTATGACCGCTGAGCAAGAGCAACAATTTGGTC	2127
Db	9546	ATTATTAATTTGTTTGAAGCACTATTTATGACCGCTGAGCAAGAGCAACAATTTGGTC	9605
Qy	2128	AACGTCCCTCTCTCCGTGGATATGCTGTAACCTGCTGCTGAATGTTTATGATACGGGA	2187
Db	9606	AACGTCCCTCTCTCCGTGGATATGCTGTAACCTGCTGCTGAATGTTTATGATACGGGA	9665
Qy	2188	CGAACAGGGAGATCCCGTCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCA	2247
Db	9666	CGAACAGGGAGATCCCGTCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCA	9725
Qy	2248	CATTTGGNAGACAGTACAGTACCTTTTCAAGCAAGTGGCAGTTCACAGAGATTTGT	2307
Db	9726	CATTTGGNAGACAGTACAGTACCTTTTCAAGCAAGTGGCAGTTCACAGAGATTTGT	9785
Qy	2308	GACCAGCGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCAGACAGTGGGT	2367
Db	9786	GACCAGCGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCAGACAGTGGGT	9845
Qy	2368	GAGTTGATCTCTTTGGGGGAGTAAATTTGAGC 2401	
Db	9846	GAGTTGATCTCTTTGGGGGAGTAAATTTGAGC 9879	

RESULT 8
AX409637
LOCUS AX409637
DEFINITION Sequence 2284 from Patent W00229103.
ACCESSION AX409637
VERSION AX409637.1 GI:21442342
KEYWORDS Homo sapiens (human)
SOURCE

AX409637 13957 bp DNA linear PAT 14-JUN-2002

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
1..13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 72.3%; Score 1735.6; DB 6; Length 13957;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
628 TGGGCAAAACATCTGTAGATGGAGACAGACCGCTGGGTCTTTTACAGACACTCATAGA 687
8216 TGGAGAGCAATTCATAAAGGGTGTAGTGGAGAGGCTGTCTTGGAGAACTCATAGA 8275
688 TTACTGCAACAGTTCCTCCCTGGACCTGGAAAGTTTCTTGCTGGCTTACAGAGCTGAA 747
8276 TTACTGCAACAGTTCCTCCCTGGACCTGGAAAGTTTCTTGCTGGCTTACAGAGCTGAA 8335
748 ACAATGCGCAATGCTCTACAGAGATCTCCCTTAAGGAAAGCTCTTAGAGAGCTCCAAG 807
8336 ACAATGCGCAATGCTCTACAGAGATCTCCCTTAAGGAAAGCTCTTAGAGAGCTCCAAG 8395
808 GGAGTAAAGAGCTGATGAAACAATGGCAAGACTCCCAAGGTGAAATGAAAGCTCAACA 867
8396 GGAGTAAAGAGCTGATGAAACAATGGCAAGACTCCCAAGGTGAAATGAAAGCTCAACA 8455
868 GATGTTTATCAACCTGGATGAAAGACCCCAAAATCTTGAGATCCCTGGAGGTTCC 927
8456 GATGTTTATCAACCTGGATGAAAGACCCCAAAATCTTGAGATCCCTGGAGGTTCC 8515
928 GATGATGCAAGTCTGTGTTACAAAGAGCTTTGGATACATGAACTTCAAGTGGAGTGAAT 987
8516 GATGATGCAAGTCTGTGTTACAAAGAGCTTTGGATACATGAACTTCAAGTGGAGTGAAT 8575
988 CGGAAAAGTCTCTCAACATTAGTCCCATTTGGAGAGCCAGTTCTGACAGCTGGAGCGT 1047
8576 CGGAAAAGTCTCTCAACATTAGTCCCATTTGGAGAGCCAGTTCTGACCACTGGAGCGT 8635
1048 CTGCACCTTTCTGAGGAGCTTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
8636 CTGCACCTTTCTGAGGAGCTTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8695
1108 CGGAGGCACTTATGGAGGAGCTTTCCAGCAGTTTCAGAGAGCAAGCATGTGATCATTAGG 1167
8696 CGGAGGCACTTATGGAGGAGCTTTCCAGCAGTTTCAGAGAGCAAGCATGTGATCATTAGG 8755
1168 GCCTTCAGAGAGGATTTGAAGATTAAGAACTGTAATCATGAGTACTCTTGAGACTGTA 1227
8756 GCCTTCAGAGAGGATTTGAAGATTAAGAACTGTAATCATGAGTACTCTTGAGACTGTA 8815
1228 CGAATATTTCTGACAGAGAGCTTTTGGAGAGCTTAGAGAACTCTTACAGAGAGCCAGA 1287
8816 CGAATATTTCTGACAGAGAGCTTTTGGAGAGCTTAGAGAACTCTTACAGAGAGCCAGA 8875
1288 GAGCTGCTCTCTGAGAGAGAGCCAGATGTCATCTGGCTTCTACGAAAGCAGGCTGAG 1347
8876 GAGCTGCTCTCTGAGAGAGAGCCAGATGTCATCTGGCTTCTACGAAAGCAGGCTGAG 8935
1348 GAGCTCAATACCTGAGTGGGAGAAATTTGAACCTGCACTCGCTGACTGGCAGAGAGAAATA 1407
8936 GAGCTCAATACCTGAGTGGGAGAAATTTGAACCTGCACTCGCTGACTGGCAGAGAGAAATA 8995

RESULT 9
AX538581
LOCUS
DEFINITION

AX538581 13957 bp DNA linear PAT 23-NOV-2002
Sequence 1 from Patent WO229056.

1408 GATGAGACCTTTCAAGACTCCAGGAATCTTCAAGAGGCCACGGATGAGTGGACCTCAAG 1467
8996 GATGAGACCTTTCAAGACTCCAGGAATCTTCAAGAGGCCACGGATGAGTGGACCTCAAG 9055
1468 CTGGCCCAAGCTGAGGTGATCAAGGGATCTGGGAGCCGCTGGGCGATCTCTCTATTGAC 1527
9056 CTGGCCCAAGCTGAGGTGATCAAGGGATCTGGGAGCCGCTGGGCGATCTCTCTATTGAC 9115
1528 TCTCTCCAAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCGCTCTGAAA 1587
9116 TCTCTCCAAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCGCTCTGAAA 9175
1588 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCGAGCTTACACATTTGGGCAATTCAGCTC 1647
9176 GAGAACGTGAGCCACGTCAATGACCTTGTCTGCGAGCTTACACATTTGGGCAATTCAGCTC 9235
1648 TCACCTGATAAATCTGAGCACTCTGGAAGACCTTGAACACCGATGGAAGCTTCTGCAAGGTG 1707
9236 TCACCTGATAAATCTGAGCACTTGTGAGACCTTGAACACCGATGGAAGCTTCTGCAAGGTG 9295
1708 GCGTTCGAGGACCGAGTCAAGGAGCTGCAATGAAGCCACAGGGAATTTGGTCCAGCATCT 1767
9296 GCGTTCGAGGACCGAGTCAAGGAGCTGCAATGAAGCCACAGGGAATTTGGTCCAGCATCT 9355
1768 CAGCACTTTCTTCCACGCTCTGTCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAA 1827
9356 CAGCACTTTCTTCCACGCTCTGTCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAA 9415
1828 GTGCCCCACTATATCAACCAACAGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA 1887
9416 GTGCCCCACTATATCAACCAACAGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA 9475
1888 GAGCTTACAGCTTTAGCTGACCTGATATATGTGAGATTTCTGAGTATATAGGATGCC 1947
9476 GAGCTTACAGCTTTAGCTGACCTGATATATGTGAGATTTCTGAGTATATAGGATGCC 9535
1948 ATGAAACTCCGAGAGCTGAGAGAGCCCTTCTGCTTGGATCTCTTGGAGCCTGTGAGTGA 2007
9536 ATGAAACTCCGAGAGCTGAGAGAGCCCTTCTGCTTGGATCTCTTGGAGCCTGTGAGTGA 9595
2008 TGTGATGCTTGGACAGCAGCAAACTCAAGGAAATGACAGCCCATGATATCTGCGAG 2067
9596 TGTGATGCTTGGACAGCAGCAAACTCAAGGAAATGACAGCCCATGATATCTGCGAG 9655
2068 ATTATTAATTTGTTGACCACTATTATGACCCCTGGAGCAAGAGCACAACAAATTTGGTC 2127
9656 ATTATTAATTTGTTGACCACTATTATGACCCCTGGAGCAAGAGCACAACAAATTTGGTC 9715
2128 AACGTCCTCTCTGCTGGATATGTGTGAACTGGCTGCTGAAATGTTTATGATACGGGA 2187
9716 AACGTCCTCTCTGCTGGATATGTGTGAACTGGCTGCTGAAATGTTTATGATACGGGA 9775
2188 CGAAGAGGAGGATCCGCTGCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA 2247
9776 CGAAGAGGAGGATCCGCTGCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA 9835
2248 CATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTGT 2307
9836 CATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTGT 9895
2308 GACCAAGGAGGCTGGGCTCTCTGCAATGATTTCTATCCAAATTCAGAGAGATTTGGT 2367
9896 GACCAAGGAGGCTGGGCTCTCTGCAATGATTTCTATCCAAATTCAGAGAGATTTGGT 9955
2368 GAAGTTGCACTCTTTGGGGGAGTAAACATTGAGC 2401
9956 GAAGTTGCACTCTTTGGGGGAGTAAACATTGAGC 9989

ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source Location/Qualifiers
1 .13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 72.3%; Score 1735.6; DB 6; Length 13957;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 628 TGGGCAACATCTGTAGATGACAGAGACCGCTGGTCTTTTACAGACACTCATAGA 687
DB 8216 TGGAGAAGCATTCATAAAGGGTGTAGTGAGCGAGAGGCTGTGGAAAGAACTCATAGA 8275
2Y 688 TTACTGCAACAGTTCCTCCCTGGACCTGGAAAGATTTCTTGGCTGTACAGAAGCTGAA 747
DB 8276 TTACTGCAACAGTTCCTCCCTGGACCTGGAAAGATTTCTTGGCTGTACAGAAGCTGAA 8335
2Y 748 ACACTGCCAATGTCTTACAGATGTACCCGTAAAGAAAGGCTCTTAGAAGACTCCAAAG 807
DB 8336 ACAACTGCCAATGTCTTACAGATGTACCCGTAAAGAAAGGCTCTTAGAAGACTCCAAAG 8395
2Y 808 GGAAGTAAAGAGCTGTAAACAAATGGCAAGACCTCCAAAGGTGAATTCGAAGCTCACACA 867
DB 8396 GGAAGTAAAGAGCTGTAAACAAATGGCAAGACCTCCAAAGGTGAATTCGAAGCTCACACA 8455
2Y 868 GATGTTTATCAACCTGGATGAAGAAACAGCCAAAAATCTCGATCTCCCTGGAAGTTCC 927
DB 8456 GATGTTTATCAACCTGGATGAAGAAACAGCCAAAAATCTCGATCTCCCTGGAAGTTCC 8515
QY 928 GATGATGCACTCTGTATCAAGACGCTTTGGATTAACATGAATTCAGAGTGGAGTGAATT 987
DB 8516 GATGATGCACTCTGTATCAAGACGCTTTGGATTAACATGAATTCAGAGTGGAGTGAATT 8575
2Y 988 CGGAAAAAGTCTCTCAACATTTAGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGGAGGCT 1047
DB 8576 CGGAAAAAGTCTCTCAACATTTAGTCCCATTTTGGAAAGCCAGTTCTGACCAAGTGGAGGCT 8635
2Y 1048 CTGACACTTCTCTGAGAGAACTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGC 1107
DB 8636 CTGACACTTCTCTGAGAGAACTTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGC 8695
2Y 1108 CGGACGACCTATTGGAGCGACTTTCCAGAGTTCCAGAGTTCAGAGCAGAACGATGTACATAGG 1167
DB 8696 CGGACGACCTATTGGAGCGACTTTCCAGAGTTCAGAGCAGAACGATGTACATAGG 8755
2Y 1168 GCCTTCAAGAGGGAATTGAAACCTTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 1227
DB 8756 GCCTTCAAGAGGGAATTGAAACCTTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 8815
2Y 1228 CGAATATTTCTGACAGCGAGCTTTGGAGGACTAGAGAACTCTTACCAGAGCCGAGA 1287
DB 8816 CGAATATTTCTGACAGCGAGCTTTGGAGGACTAGAGAACTCTTACCAGAGCCGAGA 8875
2Y 1288 GAGCTGCCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAG 1347
DB 8876 GAGCTGCCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAG 8935
2Y 1348 GAGGTCAATCTAGTGGGGAATAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407

RESULT 10

DB 8936 GAGGTCAATCTAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 8995
QY 1408 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTCAAG 1467
DB 8996 GATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGCCACGGATGAGCTGGACCTCAAG 9055
QY 1468 CTGGCCCAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGGATCTCTCTCATTTGAC 1527
DB 9056 CTGGCCCAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGGATCTCTCTCATTTGAC 9115
QY 1528 TCTCTCCAAGATCACTTCGAGAAAGTCAAGGCATTCGAGGAGAAATTCGCCCTCTGAAA 1587
DB 9116 TCTCTCCAAGATCACTTCGAGAAAGTCAAGGCATTCGAGGAGAAATTCGCCCTCTGAAA 9175
QY 1588 GAGAACGTGAGCCACGCTCAATGACCTTCTGCCAGCTTACCACTTTGGGCAATTCAGCTC 1647
DB 9176 GAGAACGTGAGCCACGCTCAATGACCTTCTGCCAGCTTACCACTTTGGGCAATTCAGCTC 9235
QY 1648 TCACCGTATAACCTCAGCACTCTGGAAGAGCTGAACACCAAGATGGAAGCTTCTGCAGGTG 1707
DB 9236 TCACCGTATAACCTCAGCACTCTGGAAGAGCTGAACACCAAGATGGAAGCTTCTGCAGGTG 9295
QY 1708 GCCGTGAGGAGACGAGTCAGGACGCTGCATGAAGCCCAAGGAGACTTTGGTCCAGCATCT 1767
DB 9296 GCCGTGAGGAGACGAGTCAGGACGCTGCATGAAGCCCAAGGAGACTTTGGTCCAGCATCT 9355
QY 1768 CAGCACTTCTTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 1827
DB 9356 CAGCACTTCTTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAA 9415
QY 1828 GTGCCCTACTATATCAACAGAGAGACTCAAAACAACTTGTCTGGAGACCATCCCAAAATGACA 1887
DB 9416 GTGCCCTACTATATCAACAGAGAGACTCAAAACAACTTGTCTGGAGACCATCCCAAAATGACA 9475
QY 1888 GAGCTCTACCACTCTTTAGCTGACCTGAATAATGTCTCAGATTCTCAGCTTATAGACTGCC 1947
DB 9476 GAGCTCTACCACTCTTTAGCTGACCTGAATAATGTCTCAGATTCTCAGCTTATAGACTGCC 9535
QY 1948 ATGAAATCTCGAAGACTTGCAGAGAGGCGCTTTTGGTGTGATCTCTTGGCCCTGTGAGTGCA 2007
DB 9536 ATGAAATCTCGAAGACTTGCAGAGAGGCGCTTTTGGTGTGATCTCTTGGCCCTGTGAGTGCA 9595
QY 2008 TGTGATGCTTGGACAGCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCTCTGAC 2067
DB 9596 TGTGATGCTTGGACAGCACAACCTCAAGCAAAATGACCAAGCCCATGGATATCTCTGAC 9655
QY 2068 ATTATTAAATTTGTTGACCACTATTATGACCGCTGGAGCAAGAGCACAACAATTTGGTC 2127
DB 9656 ATTATTAAATTTGTTGACCACTATTATGACCGCTGGAGCAAGAGCACAACAATTTGGTC 9715
QY 2128 AACGTCTCTCTGCGTGGATATGTCTGAATCTGCTGCTGCTGAATGTTTATGATACGGGA 2187
DB 9716 AACGTCTCTCTGCGTGGATATGTCTGAATCTGCTGCTGCTGAATGTTTATGATACGGGA 9775
QY 2188 CGAACAGGAGGATCCGCTGCTCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 2247
DB 9776 CGAACAGGAGGATCCGCTGCTCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 9835
QY 2248 CATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGT 2307
DB 9836 CATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGT 9895
QY 2308 GACCAGCGAGGCTGGGCGCTCTTCTGATGATTTCTATCCAAATTCACAGACAGTTGGGT 2367
DB 9896 GACCAGCGAGGCTGGGCGCTCTTCTGATGATTTCTATCCAAATTCACAGACAGTTGGGT 9955
QY 2368 GAAGTTGCATCTTTTGGGGCAGTAACTTGGC 2401
DB 9956 GAAGTTGCATCTTTTGGGGCAGTAACTTGGC 9989

HUMDYS
LOCUS 13957 bp mRNA linear PRI 25-MAY-2000
Homo sapiens dystrophin (DMD) mRNA, complete cds.
ACCESSION M18533 M17154 M18026 M20250
VERSION M18533.1 GI:181856
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1699)
Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C. and
Kunkel, L.M.
Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
preliminary genomic organization of the DMD gene in normal and
affected individuals
Cell 50 (3), 509-517 (1987)
JOURNAL 87273512
MEDLINE 3607877
PUBMED
REFERENCE 2 (bases 1678 to 3830)
Hoffman, E.P., Monaco, A.P., Feener, C.C. and Kunkel, L.M.
Conservation of the Duchenne muscular dystrophy gene in mice and
humans
Science 238 (4825), 347-350 (1987)
JOURNAL 88018015
MEDLINE 3659917
PUBMED
REFERENCE 3 (bases 1 to 13957)
Koenig, M., Monaco, A.P. and Kunkel, L.M.
The complete sequence of dystrophin predicts a rod-shaped
cytoskeletal protein
Cell 53 (2), 219-226 (1988)
JOURNAL 88194521
MEDLINE 3282674
PUBMED
COMMENT
On May 25, 2000 this sequence version replaced gi:340693.
Draft entry and computer-readable sequence kindly provided by
M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
determined by the size of the deleted DNA segment. Deletions found
in different patients were from positions 302-2200, 473-1168,
1691-1810, and 1169-3011.
FEATURES
source
1..13957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"
1..13957
/gene="DMD"
209..11266
/gene="DMD"
/codon_start=1
/product="dystrophin"
/protein_id="AAA53189.1"
/db_xref="GI:181857"
/translation="MLWMEVEDCYBERDVQKFTTKVNAQFSKFGKHILNLFSDL
QDGRLLDLGLTGOKLPKXGSTRVHALNNKALVLONNVYDLVNIQSDIVDG
NHKLTGLIWNILHQVKNVKNIMAGLQNTSNKILLSWRQSTRYPQVNIPT
TSWGLALNHLHSRDPFOWNSVYCCQSATQRLEAFNARIQLGIEKLLDPEDV
DTTPDKSILNHYTSLSFQVLPQVQVSEIAQIEVEMLPKPPKVBEEHQLHQMHSYQ
QITVLAQGYERTSSPKPFKSYATQAAVYVTSPTSPSPPSQHLAPEDKSFQSL
MESEVNLDRYQTALESVLSWLSAEDTLQAQGEISNDVEVVKDQFTHGYNMGLTAH
QGRVGNILGSKLIGTKLSDETEVEQMNLSNRWECILRVASMEKQSNLHRLM
DLNCKLKNLMDLTKTERTKMREEPGLPDLKQVQOQKVLQDLBQEOVRVN
SLTHMVVVDSSGDHATAALBEQLKVLGDRWANICRWTEBHWLVLLQILLKWLTE
EQCLFASLSEKEDAVNKLHTTGCDQDNEMLSLQKJAVLRKADLEKKQSGMKGLSLK
QCLLTKNKSVTQTEAWLDFARCDNLVQLKJSAQISQAVTTQPSLTQTVN
ETVTTVTRFRLHQHQAQLPPPPKQKQITVDSIRKRLDVIDITELHWNITRSEA
VLOSPPFAIFKEGNFSDLEKVNATERKAEKFKRLQDASRAQALVEQMNVEGVA
DSIKQASEOLNSWIEFQCLLSERLNWLEVOYNNIAFYNOQLQEQMTTAAENWLKIQ
PTTPSPPTAKSKICKDEVNLSGLQPIERLKIQSIALKEKQGGPWFDAQVAF

TNHFQVFSVDQAREKELOTIPDTLPMRYQETMSAIRTWVQOOSTKLSIPOLSVTDY
EIMEQGEIQAQSSIQEOQSSLYLVLTIVKEMSKAPSEISRKYQSPFEIEGWRK
KLSQVLEHCQKLEEQNKLRKIQNHIOQLKWAEDVDYFLKEEALGDSEILKKQL
KQCRLLVSDIQTIQPSLNSVNEGQKIKNEAEFEFASRLETELKELNTQNDWCQOY
ARKEALGGLEKTVTSQKDLSEHWMWTQABEYLERDEFKTPDLSQXKAVEMKRAK
EWAQCKEAKKLTSTESVNSVIAQAPPAQBAALKELETTITTYOWICTLNGKCKTLE
NQRILQAOTLDGGMDELINLELETNFSNRWELHEEAVRQKLSQSAQSETSKS
LHLIQSLTFIDKQAAIADKVAQMPQEAOKIOSDLTSHIEISLEEMKHNQGEKA
AQVLSQIDVAQKLDQVSMKPFLOKPAFELRLOESKWLIDEVWMLPALETKEVE
QEVVQSLNHCNLYKLSLVKSEVEMVITKGRQIVQKQTEPNKEDRVTALKLHY
NEGAKVTERKQKLEKALSKLRKWRKEMNVLTELAATDMEDLTKRAVEMPSNLSE
VANGATKEIEKQKHLKSIITEVGAITVKLGKETLVELDKLSLNSNWIATISRAE
IRWNLLEYQKMETFDONVDHITKWIIOADTLDESEKKKQKQEDVLKRLKLEND
EINKVSTQRQAAANLMDGDKLVEPQISLHNRFASISHRITKVASIPLEKLE
QFNSDIQKLEPLEAEIOGVNLKEDFNKMDNEGTVEKLLQDGNLQRIIDEBK
KRBEIKIQOOLKONALKLSORRKALEISHOWYOYKQADDLKCLDDIBEKL
ASLPEPRDERKIKELDEIQKKEELNAVROAGLSEGGAMAVPTQIOISKRWE
TESFAQFRKINFAQIHVREETMWNTEDELEISYFSTVLTETHTHVSQALLEVQ
LNAAPDLCAKDFDLFKQESLKNIKDSIQQSGRIDIIHSKKTALQALQATPVERVKL
QEALSQDFQWKNVKNYKDRQGRFDRSVKWRPHYDIKFNQWTEAELFLKTCQI
PENWEHAKYKWLKELQDGIQGTQVTVTLNATGEEIIQOSSKTASIIQELKLSNL
RWQEVCKQSDRKLREKQNLSEFORDLNEFVLMEADNIASIPLEGKQOOLKE
KLQVKLLVEELPLQIGILKQNLNETGSPVLVSAPISEBQDKLENKLTQNLQWIKVS
RALPEKQGEIEAQIKDLQLEKLELRLQNLHLLWLSPIRQLEIYNQPOGEPD
VQSTETAVQAKQPDVEEILSKQHLKEXRQATQPVKRLKEDLSSEKAVNRLLEURA
KQFDLAPGLTTIGASPTQTVTLVTPVTKETAISKLEPSSMLKVPALADFNRAWT
ELTDWLSLDQVIKSQRVVMQDLEENIMIIOKATMQDLQORRPOLELITAEAGVL
NKTSGOARTIITDRIERIQNDQVEHQLQNRQQLNEMKLDSTOWLEBAKEAOVL
GOARAKLESWKEGPTVDAIKKITERKOLAKDQWOTNVDVANDLAKLRLDYSAD
DTRKHMVITENINASWSTHKVSREAALETHRLQQLPDLLEKFLAWLTAETAT
NVLODATRKERLLEDKSGYKELMKQWQDQLEIAHDTYHNLDSQKLSLEGSD
DAVLQRLDNNFKNWSELKSNLNRKSHLEASSDQKRLHLSLQELLVWQLKDBEL
SRQAGIGGPPFAVQKQNDVHRAPRELKTEKPMVINSTLETIRIFTEPLEGKEUQY
EPRELPEERAQNVTLRKQAEVNTWEKLNLSHADQKXIDELEQLQLEQATD
ELDLKQREVIKGSQVGDLLIDSLQHLKVKALRGCIAPLEKNSVHNDLARQL
TTLGIOLSPVNLSTLDMRLKVLQVAVEDVRQLHEARDFGAPSOHFLTSVQGP
WETAIISPKVPYVNIHETOTQDHPKMTVELYCSLADLNNVRFSVARTAKLRLQKA
LCIDLLSAAQDALDQNLKONDOPMDILOIINCLTTIYDRLEOHNHNVNPLCVD
KCLNWLNVDTQRTGRIIRVLSFKGILISICAKAHLDEKRYLFKQVASTGSGDQRL
GLLHDSIQIPQLQGEVSGSGNSBPFSVCFQFANNKPEIEAALFLDMWRLEPQSM
VWLPVLRVAAATAXHQAKNICEPIIGFRYSLKHFNVDICOCFSFGVRKXGH
KMHYPVVEYCTTSGEDVDRFAKLVNFKTRKRYFAKPRMGYLFPQVTLQEDNMET
PVTLLNFWPVSAPASSPOLSHDDTHSRIEHSVLAEMENGSYNDISNPESID
DEHLIHCYQSLQNSQPSLQSPSPAIIISLESEERBELERILADLEENLNQAEY
DRLKQHEHKGSLPLSPPEMPTSPQSPRDAELIAEAKLQKHGKRLERARQOILBDH
NKQESQLHRLQLLEQPOAEKYNNGITVSSPSLQSDSSQPMLLRVRVQTSQSDSM
GEEEDLLSPQDSTGTGLEEVQELNNFSPSRGNTPEKPMRBDTM"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 72.3%; Score 1735, 6; DB 9; Length 13957;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 628 TGGCAACATCTGTAGATGACAGAGACCGCTGGTGTCTTTTACAGACACTCATA 687
DB 8216 TGGAGAAGCATTCATAAAGGGTGAGTGACGAGAGCGTCTTTTGGAGAAACTCATA 8275
QY 688 TTACTGCAAGTCCTCCCTGGACCTGGAAGAGTTTCTTGCCTGGCTTACAGAGCTGAA 747
DB 8276 TTACTGCAAGTCCTCCCTGGACCTGGAAGAGTTTCTTGCCTGGCTTACAGAGCTGAA 8335
QY 748 ACACTCCCAATGTCTTACAGAGTGCTACCCGTAAGAAAGGCTCCTAGAAAGACTCCAAG 807
DB 8336 ACACTCCCAATGTCTTACAGAGTGCTACCCGTAAGAAAGGCTCCTAGAAAGACTCCAAG 8395
QY 808 GGAGTAAAGAGCTGATGAAACAAATGCGAGACCTCAAGGTGAATTGAAGTCAACACA 867
DB 8396 GGAGTAAAGAGCTGATGAAACAAATGCGAGACCTCAAGGTGAATTGAAGTCAACACA 8455
QY 868 GATCTTTATCAACACTGGATGAAAAACAGCAAAAAATCTTGAGATCCCTCGAAGGTTC 927

Db 8456 GATGTTTATCACAACTGGATGAAAAAGCCAAAAAATCTTGAGATCCCTGGAGGTTCC 8515
QY 928 GATGATGAGTCCTGTTTACAAAGACGTTTGGATTAACATGAACCTTCAAGTCGAGTGAACCT 987
Db 8516 GATGATGAGTCCTGTTTACAAAGACGTTTGGATTAACATGAACCTTCAAGTCGAGTGAACCT 8575
QY 988 CGGAAAAAGTCTCTCAACATTTAGTGTCCATTTTGAAGCCAGTGTCTGACCAAGTCGAGTCG 1047
Db 8576 CGGAAAAAGTCTCTCAACATTTAGTGTCCATTTTGAAGCCAGTGTCTGACCAAGTCGAGTCG 8635
QY 1048 CTGCACCTTTCTCTGCAGGAACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGC 1107
Db 8636 CTGCACCTTTCTCTGCAGGAACTTCTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGC 8695
QY 1108 CGGAGAGCACTTATGAGAGGCACTTCCAGCAGTTCAGAGCAGACAGATGATGATAGG 1167
Db 8696 CGGAGAGCACTTATGAGAGGCACTTCCAGCAGTTCAGAGCAGACAGATGATGATAGG 8755
QY 1168 GCCTTCAAGAGGGAATTTGAAACCTTAAAGAACCTGTATCATGATGACTCTTGTAGACTGTA 1227
Db 8756 GCCTTCAAGAGGGAATTTGAAACCTTAAAGAACCTGTATCATGATGACTCTTGTAGACTGTA 8815
QY 1228 CGAATATTTCTGACAGAGAGCCCTTTTGGAAAGACTAGAGAAACTCTACAGAGGCCAG 1287
Db 8816 CGAATATTTCTGACAGAGAGCCCTTTTGGAAAGACTAGAGAAACTCTACAGAGGCCAG 8875
QY 1288 GAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCTACGAAGCAGGCTGAG 1347
Db 8876 GAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCTACGAAGCAGGCTGAG 8935
QY 1348 GAGTCAATATCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407
Db 8936 GAGTCAATATCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 8995
QY 1408 GATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGGCCACGATGAGCTGACCTCAAG 1467
Db 8996 GATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGGCCACGATGAGCTGACCTCAAG 9055
QY 1468 CTGCGCCAGCTGAGGTGATCAAGGGGATCTGGCAGCCCTGGCGCATCTCTCATTTGAC 1527
Db 9056 CTGCGCCAGCTGAGGTGATCAAGGGATCTGGCAGCCCTGGCGCATCTCTCATTTGAC 9115
QY 1528 TCTCTCCAGATCACCTCGAAGAAAGTCAAGGCACTTCGAGAGAAATTCGGCTCTGAAA 1587
Db 9116 TCTCTCCAGATCACCTCGAAGAAAGTCAAGGCACTTCGAGAGAAATTCGGCTCTGAAA 9175
QY 1588 GAGAACTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCACTTTGGCATTGAGTC 1647
Db 9176 GAGAACTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCACTTTGGCATTGAGTC 9235
QY 1648 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTG 1707
Db 9236 TCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAGGTG 9295
QY 1708 GCGTTCAGAGACCGAGTCAGGAGCTGCATGAAGCCACAGGACCTTTGGTCAGCATCT 1767
Db 9296 GCGTTCAGAGACCGAGTCAGGAGCTGCATGAAGCCACAGGACCTTTGGTCAGCATCT 9355
QY 1768 CAGCACTTTCTTTCCAGCTGTCTCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAA 1827
Db 9356 CAGCACTTTCTTTCCAGCTGTCTCCAGGGTCCCTGGGAGAGCCATCTCGCCAAACAAA 9415
QY 1828 GTGCCCTTACTATATCAACACAGAGACTCAAAACATTTGTGTGGAGCCATCCCAAAATGACA 1887
Db 9416 GTGCCCTTACTATATCAACACAGAGACTCAAAACATTTGTGTGGAGCCATCCCAAAATGACA 9475
QY 1888 GAGCTTACCACTGCTTTAGCTGACCTGGAATAATGTCAGATTTCTCAGCTTATAGGACTGCC 1947
Db 9476 GAGCTTACCACTGCTTTAGCTGACCTGGAATAATGTCAGATTTCTCAGCTTATAGGACTGCC 9535
QY 1948 ATGAACTCCGAAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCA 2007
Db 9536 ATGAACTCCGAAAGACTGCAAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCA 9595

QY 2008 TGTGATGCTTTGGACCCAGCACAACTCAAGCAAATGACCCAGCCATGATATCTCTGCAG 2067
Db 9596 TGTGATGCTTTGGACCCAGCACAACTCAAGCAAATGACCCAGCCATGATATCTCTGCAG 9655
QY 2068 ATTATTAATTTGTTGACCACTATTATGACCGCTGGAGAGGACCAACAATTTGGTC 2127
Db 9656 ATTATTAATTTGTTGACCACTATTATGACCGCTGGAGAGGACCAACAATTTGGTC 9715
QY 2128 AACGTCCCTCTCTGCTGGTGAATGCTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGA 2187
Db 9716 AACGTCCCTCTCTGCTGGTGAATGCTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGA 9775
QY 2188 CGAACAGGAGATCCGTCCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA 2247
Db 9776 CGAACAGGAGATCCGTCCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA 9835
QY 2248 CATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 2307
Db 9836 CATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGT 9895
QY 2308 GACCAAGCAGGCTGGGCTCTCTCTGATCATTTTCAAAATTCAAAGACAGTTGGGT 2367
Db 9896 GACCAAGCAGGCTGGGCTCTCTCTGATCATTTTCAAAATTCAAAGACAGTTGGGT 9955
QY 2368 GAAGTTGCATCCTTTGGGGCAGTAACATTGAGC 2401
Db 9956 GAAGTTGCATCCTTTGGGGCAGTAACATTGAGC 9989

RESULT 11

AX538620 LOCUS AX538620 5339 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source 1..5339
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"

BASE COUNT 1638 a 1191 c 1187 g 1323 t

ORIGIN

Query Match 71.8%; Score 1725; DB 6; Length 5339;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1721; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 681 TCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTGCTGGCTTACAGA 740
Db 1546 TCATAGATTACTGCAACAGTTCCCTGGACCTGGAAAGTTTCTGCTGGCTTACAGA 1605
QY 741 AGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTTCCTAGAGA 800
Db 1606 AGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTTCCTAGAGA 1665
QY 801 CTCGAAGGGATTAAGAGCTGATGAACAATGGCAAGACTCCAGGTGAATTAAGC 860
Db 1666 CTCGAAGGGATTAAGAGCTGATGAACAATGGCAAGACTCCAGGTGAATTAAGC 1725
QY 861 TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCTTGAGATCCCTGGA 920
Db 1726 TCACACAGATGTTTATCACAACCTGGATGAAAAACAGCCAAAAATCTTGAGATCCCTGGA 1785

921 AGTTCCGATGATGAGTCTCTGTTACAAAGACGTTTGGATACATGAATCAAGTGGAG 980
1786 AGGTTCCGATGATGAGTCTCTGTTACAAAGACGTTTGGATACATGAATCAAGTGGAG 1845
981 TGAATCTCGGAAAAAGTCTCTCAACATAGTCTCCATTTGGAAGCCAGTCTTGACAGTG 1040
1846 TGAATCTCGGAAAAAGTCTCTCAACATAGTCTCCATTTGGAAGCCAGTCTTGACAGTG 1905
1041 GAAGCTCTGCACTTTCTCTGAGGAACTCTCTGTTGGTGTGCTTACAGCTGAAAGATGATGA 1100
1906 GAAGCTCTGCACTTTCTCTGAGGAACTCTCTGTTGGTGTGCTTACAGCTGAAAGATGATGA 1965
1101 ATTAAGCCGCGAGGCACTTATTTGAGGCACTTTCCAGCAGTTTCAAGACGAAAGATGT 1160
1966 ATTAAGCCGCGAGGCACTTATTTGAGGCACTTTCCAGCAGTTTCAAGACGAAAGATGT 2025
1161 ACATAGGCGCTTCAGAGGAACTTGAAGAACTTGAAGAACTTGAAGAACTTGAAGAACTTGA 1220
2026 ACATAGGCGCTTCAGAGGAACTTGAAGAACTTGAAGAACTTGAAGAACTTGAAGAACTTGA 2085
1221 GACTGTACGAATATTTCTGACAGAGCAGCTTTTGAAGGACTTAGAGAACTTACACGGA 1280
2086 GACTGTACGAATATTTCTGACAGAGCAGCTTTTGAAGGACTTAGAGAACTTACACGGA 2145
1281 GCCAGAGAGTGCCTCTCTGAGAGAGAGCCAGAGATGTCTCGGCTTCTACGAAAGCA 1340
2146 GCCAGAGAGTGCCTCTCTGAGAGAGAGCCAGAGATGTCTCGGCTTCTACGAAAGCA 2205
1341 GGCTGAGGAGGTCAATPACTGAGTGGGAAATTTGAACCTGCACTCGCTGACTGGCAGAG 1400
2206 GGCTGAGGAGGTCAATPACTGAGTGGGAAATTTGAACCTGCACTCGCTGACTGGCAGAG 2265
1401 AAAAAATAGTAGAGACCTTTGAAAGACTTCAGGAACCTTCAAGAGCCAGAGATGAGCTGA 1460
2266 AAAAAATAGTAGAGACCTTTGAAAGACTTCAGGAACCTTCAAGAGCCAGAGATGAGCTGA 2325
1461 CCTCAAGCTGCGCAAGCTGAGTGATCAAGGATCTCTGCGAGCCGCTGGGCGATCTCT 1520
2326 CCTCAAGCTGCGCAAGCTGAGTGATCAAGGATCTCTGCGAGCCGCTGGGCGATCTCT 2385
1521 CATTTGACTCTCTCCAGATCACTCTGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGCGCC 1580
2386 CATTTGACTCTCTCCAGATCACTCTGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGCGCC 2445
1581 TCTGAAAGAGAACTGAGCCAGCTCAATGACCTTCTGCGCAGCTTACCACCTTTGGCGAT 1640
2446 TCTGAAAGAGAACTGAGCCAGCTCAATGACCTTCTGCGCAGCTTACCACCTTTGGCGAT 2505
1641 TCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAAGATGGAAGCTTCT 1700
2506 TCAGCTCTCAACCGTATAACCTCAGCACTCTGGAAGACCTTGAACACCAAGATGGAAGCTTCT 2565
1701 GCAGTGGCGCTGAGGACCGAGTCAAGGAGCTGATGAAGCCCAAGGCACTTTGGTCC 1760
2566 GCAGTGGCGCTGAGGACCGAGTCAAGGAGCTGATGAAGCCCAAGGCACTTTGGTCC 2625
1761 AGCATCTCAGCACTTTCTTCCAGCTCTGTCCAGGCTTCCAGGAGAGCCATCTCGCC 1820
2626 AGCATCTCAGCACTTTCTTCCAGCTCTGTCCAGGCTTCCAGGAGAGCCATCTCGCC 2685
1821 AAACAAAGTCCCTTACTATATCAACCAAGAGTCAAAACAACTTGTGTGGGACATCCCAA 1880
2686 AAACAAAGTCCCTTACTATATCAACCAAGAGTCAAAACAACTTGTGTGGGACATCCCAA 2745
1881 AATGACAGAGCTTACAGCTTTTGTGACCTTGAATATGTCTGAGATCTTCTGAGCTTATAG 1940
2746 AATGACAGAGCTTACAGCTTTTGTGACCTTGAATATGTCTGAGATCTTCTGAGCTTATAG 2805
1941 GACTGCCATGAACTCCGAGAGCTGAGAGGCGCTTTGCTTGGATCTCTTGTAGGCTGTG 2000
2806 GACTGCCATGAACTCCGAGAGCTGAGAGGCGCTTTGCTTGGATCTCTTGTAGGCTGTG 2865

2001 AGCTGCATGTGATGCTTGGACAGCACAACTCAAGCAAAATGACAGCCCATGGATAT 2060
2866 AGCTGCATGTGATGCTTGGACAGCACAACTCAAGCAAAATGACAGCCCATGGATAT 2925
2061 CCTGCAGATATTAAATGTTTGAACCACTATTTATGACCGCTGGAGCAGACACAA 2120
2926 CCTGCAGATATTAAATGTTTGAACCACTATTTATGACCGCTGGAGCAGACACAA 2985
2121 TTTGGTCAACGTCCTCTCTGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGA 2180
2986 TTTGGTCAACGTCCTCTCTGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGA 3045
2181 TAGCGAGCAAGAGGAGGATCGTCTGCTCTTTTAAACTGGCATCATTTCCCTGTG 2240
3046 TAGCGAGCAAGAGGAGGATCGTCTGCTCTTTTAAACTGGCATCATTTCCCTGTG 3105
2241 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAG 2300
3106 TAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAG 3165
2301 ATTGTGTGACCGAGCGAGGCTGGGCTCTCTCTGCTGATGATTTCTATCCAAATTTCAAGACA 2360
3166 ATTGTGTGACCGAGCGAGGCTGGGCTCTCTCTGCTGATGATTTCTATCCAAATTTCAAGACA 3225
2361 GTTGGGTGAAGTTGCTATCCTTTGGGGGCGAGTAACTTGAGC 2401
3226 GTTGGGTGAAGTTGCTATCCTTTGGGGGCGAGTAACTTGAGC 3266

RESULT 12
AR220819
LOCUS AR220819 13977 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 60 from patent US 6426186.
ACCESSION AR220819
VERSION AR220819.1 GI:23327696
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmut,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES
Location/Qualifiers
1..13977
source /organism="unknown"
BASE COUNT 4596 a 2765 c 3120 g 3453 t 43 others
ORIGIN

Query Match 71.8%; Score 1724.6; DB 6; Length 13977;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 628 TGGCAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTACAAGACATCTATAGA 687
DB 8216 TGGAGAGCACTCATAAAGGGTGAAGCGAGAGGCTGCTTTTGAAGAAATCTATAGA 8275
QY 688 TTACTGCAACAGTTTCCCTCTGGACCTGGAAAAGTTTCTTCCCTGGCTTACAGAAGCTGAA 747
DB 8276 TTACTGCAACAGTTTCCCTCTGGACCTGGAAAAGTTTCTTCCCTGGCTTACAGAAGCTGAA 8335
QY 748 ACACTGCCAATGTCTTACAGGATGTCTACCCGTAAAGAAAGGCTCTTAGNAGACTCCAAG 807
DB 8336 ACACTGCCAATGTCTTACAGGATGTCTACCCGTAAAGAAAGGCTCTTAGNAGACTCCAAG 8395
QY 808 GGAGTAAAGAGCTGTGAAACAATGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 867
DB 8396 GGAGTAAAGAGCTGTGAAACAATGCAAGACCTCCAAGGTGAAATTTGAAGCTCACACA 8455
QY 868 GATGTTTATCAACACTCTGATGAAAACAGCCAAATAATCTCTGAGATCCCTTGAAGGTTCC 927
DB 8456 GATGTTTATCAACACTCTGATGAAAACAGCCAAATAATCTCTGAGATCCCTTGAAGGTTCC 8515

YDIMEORLQALQALQLOEQNGNLNLTSTTVKEMSKAPLSDISIRKYQSEFEBIEGR
WKLSSQVHEHCQKLEQWAKLRLQNHKTLKWLTEVDVLEKEWALGDSILKE
QLKQKLLVNDLQITQPSLNSVNEGAQKQKNEAEPFARLETELRELNTQNDYACRQ
VYARKEALGGDLKDTLQKLDUSEHEWMTQABEELRDEFEYKTPDQTAWEKMR
AKEEQAQKAKVLLTESVNSVIAQAPPAQKALKELDITLQWLTQRLNCKST
LEBEVACWHELLEYLEKANKWLSVEVLLKTTENTSGGAETIAEVLSLENLQHSDE
NPQNRILAQITLTDGWMDELINLETFNSRWELHEEAVRQKLRQKLSQSAQIE
KSLHLOBSLSDIKQALVADKYDAQMPQOAKIOSDLTSHBISLEENKKNQOG
EYQVRVLSQIDVQAKKLDVSKFRLFOKPAFQLOESKWLDEVKWHLPALETKS
VQEYVQSQNGNCVNLKSLSEVSEVENVIKGTQIQVKQKOTENFKELDERVTAHL
HYNELGAKVTERQLEKCLKLSRMRKEMNALTLEWLAQNTMELTQRAVEGMSNLD
SEVAMGATQKIEKQKLVTEVEGALKTVLQKXELVEDKXELKSLNSLNSWIAVSR
ABEMLLLELYQKMETFQNDYITNWIQADALLDESEKVKPOKQEDILKRLKAM
NDRPVDGTRODAANLNRQGHCRKVVEPKISELNRFAISRIRIKGTAKSLPKE
LEQFNSDIQKLEPFEABEQVNLKEDEPNKMSDENEGTVKELLQKQNLQORITD
EKREBEIKIQQLOTLQKHALKDLRSQREKKALEISHOWYOYKROADLLKCLDDIEK
KLASPEPDERKIKEDIRELOKKEELNVRQAEGLSEDAANVAEPTQIQLSKRW
RELSEKPAFRRLNFQIHLVHEESVAMTMDPLEISVFPSTYILTEITHVSQALSEV
EELNAPDLCAQDFEDFKQESLKNIDSLQIQSGRIDIHKKTAALHSAFPAERA
KLOEALSRDLDFQWRVNNYKQGRFDRSVKRRPHYDMKILNQMITEAEQFLKKT
QIPENNEHAKYKWLXELQDQIGQOSVVRVNLATGEBIIQOOSKTDASILQEKLSL
NLRQVQVKQLAERKRLBEOKNLSSEFORDVNEFVLEADNVANTPLRPGNEQOL
KEKLEQVLLAEELPLRQKILQKLNQETGTVLVSAPLSPEEQDKLENKIKOTNLQWIK
VSRNLEKQEEBEAHVKDQLEQELNHLILLWLSPIRQLEIYNGPOTGTFDKEIE
VAVQAKQPDVEGLSGQHLXKEKPAQKAKKLDLSSDMKVVTQLLOELRAKQPGP
APGLTTRAPPSPQVTVLQPAVTKETAKSLEPSSLLLEVPALADFNRAWELTDW
LSLLDRVKSQRVWVGLDEDINEMIKOKATLOLQORPOLEELITAOONLKNKTSN
QOEARTITRERIQRQDEVEHQNRRLQOLTEMLKSTOWLEAKSEAEQVLQARA
KLESWEAPYVDAIQKTTETKOLAKDLQWQINVDVANDLAKLLRDYSADDTKRV
HMTENINASASIHRELSEKRAALBEETHLLQOPLDLEKELWLEAETANVLQD
ATHKRLLEDSSGRKMLQWQLOQSEIEAHTDIYHNDENGQKVLRSLESGDDAALL
QRKLDNNFKSRLKKSINLSHLEASSDQWRLHLSLQELWLLQKQDDELSSQAP
IGDFTFAVQNDVHAFREKLEKPEVIMSTLETRIFLPEQLEGLKELYQREPL
PPEERQNTVRLRKQAEVNTQWELKNVNSADWQKIDLEALQELQELQAELEDELK
LROAEVITKGSWQDGLLDSLQDHLKVKALRGHITPLKENVSVYNDLAROLATLGI
QLSPYNLNTDLENTKLOVAIEDRIQLHEARDFGPSQHFSLTSVOGMPERAI
SPNKVYVYNHETQTCNDHPKMTLEYQSLADLNNVRPSAVETAKVLARLOKALCLD
LSLSACADLQKHLKQNDQMDILQVINCITTIIDRLQENHNNLNVPLQVDMCLNW
LNVYDTRGTRIRVLISFTYLSICAKHLEDKRYLFKQVASSTFGCDQRLGLLHL
DSQIPRQLEGAFAVGSNIPEVRSFCFQANNKEIEEALFLDWRLPEQSMWLPV
LHRVAAAEAKHQAKNICKEQPIIGFVRSILKFNHNDICQSCFFSGVAKGHXHP
MVEYCTPTTSGEVRDFKVLKNKTRKRYFAKHPRMGVLPVQTVLEGNMNETPVLLI
NFWPVSAPASSPOLSHDDTHSRIEYHVASRLKMNNSGVLNDSISPNESIDDBHLI
IQHWASLQCESPLSQPRAPQAILLSLEEERLELERILADLEGRNRLQAEYDRKLO
QHEGLGSLFSPSEMPSTPSQPRADLIEAKLLRQKHLERMLQLEHNDKQLE
SQHLRQLLEQPOAKVNGTIVSPSTLQORSOSSQWLLRUVVGSQTSSESMBEEDL
LSPQDSTSTGLEEVMELNHSFPSSRRNTPCKPMREDTM"
11333. .>13887
11789. .11828
/rpt_type=tandem
/rpt_unit=ca
3'UTR repeat_region 4455 a 2906 c 3164 g 3362 t

BASE COUNT 4455 a 2906 c 3164 g 3362 t
ORIGIN

Query Match 64.2%; Score 1540.4; DB 4; Length 13887;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1628; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 628 TGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAGACACTATAGA 687
Db 8282 TGGCAAGCATCCATAAAGATTGAGTGAGCGAGAGCTGCTCTGGAAGAAACCCACAGA 8341

Qy 688 TTACTGCAACAGTTCCTCCCTGGACCTGGAAAGTTCTTTCCTGGCTTACAGAGCTGAA 747
Db 8342 TTACTGCAACAGTTCCTCCCTGGACCTGGAGAGTTCTTTCCTGGCTTACAGAGCTGAA 8401

Qy 748 ACACTGCCAATGCTTACAGGATGCTACCCGTAAGAAAGGCTCTCTAGAGACTCCAAG 807
Db 8402 ACACTGCCAAGCTCTGCGAGGATGCCACCCATAAGAAAGGCTCTCTAGAGATTCCAAG 8461

Qy 808 GGAGTAAAGAGCTGATGAACAAATGCGAAGACTCCAGAGCTCCAGGTGAATTTGAAGCTCACACA 867
Db 8462 GGAGTAAAGAGCTGATGAACAAATGCGAAGACTCCAGAGAAATTCGAAGCTCACACA 8521

868 GATGTTTATCACAACCTGGATGAAAAACGCAAAAAATCCTGAGATCCTGGAAGCTTCC 927
8522 GATATCTATCACAACCTGGACGAAAAATGCGCAAAAAAGTCTGAGATCCTCGAAGTCT 8581

Qy 928 GATGATGAGTCTCTGTTCACAAAGACCTTTTGGATAACATGAATTTCAAGTGGAGTGAACCT 987
Db 8582 GAGCATGAGCCTTTGTTCGAAAGACGTTTGGATPAACATGAATTTCAAGTGGAGCAGAACT 8641

Qy 988 CGGAAAAAGTCTCTCAACATTTAGTGTCCCTTTTGGAAAGCCAGTCTTCCACAGTGGAAAGCGT 1047
Db 8642 CGGAAAAAGTCTCTCAACATTTAGTGTCTTCTGGAAGCCAGTCTTCCACAGTGGAAAGCGT 8701

Qy 1048 CTGCACCTTTCTCTGCAAGAACTTTCTGGTGTGCTACAGCTGAAAGATGATGAATTAAGC 1107
Db 8702 CTGCACCTTTCTCTGCAAGAACTTTCTGGTGTGCTCCAGCTGAAAGATGATGAATTAAGC 8761

Qy 1108 CGGAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCCAGAGCAGACGATGTGATACATAGG 1167
Db 8762 CGGAGGACCTATTGGAGGCGACTTTCCAGCAGGCTGCAAGAGCAGAAATGATGTACACAGG 8821

Qy 1168 GCCTTCAAGAGGAAATTTGAAAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 1227
Db 8822 GCCTTCAAGAGGAAATTTGAAAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTA 8881

Qy 1228 CGAATATTTCTGACAGGAGCCTTTTGGAAAGGACTAGAGAACTCTACAGGAGCCACAGA 1287
Db 8882 CGAATATTTCTGACAGGAGCCTTTTGGAAAGGACTAGAGAACTCTACAGGAGCCACAGA 8941

Qy 1288 GAGCTGCTCTCTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCTACGAAAGCAGGCTGAG 1347
Db 8942 GAGCTGCTCTCTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCTACGAAAGCAGGCTGAG 9001

Qy 1348 GAGTCAATACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATA 1407
Db 9002 GAGTCAATACTGAGTGGGAAAAATTTGAACCTGCACTCTGCGACTCTGCGAGACTGGCAGAAAAATA 9061

Qy 1408 GATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGCCAGCGATGAGCTGGACCTCAAG 1467
Db 9062 GATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGCTTCAAGAGCAGACAGATGAGCTGAA 9121

Qy 1468 CTGCGCCAACTGAGGTGATCAAGGGATCTCGACGCCCGTGGCGGATCTCTCATTTGAC 1527
Db 9122 CTACGTGAGGACGAGGTGATCAAGGGATCTCGACGCCCTGTGGGTGACCTCTCATTTGAC 9181

Qy 1528 TCTCTCAAAGTACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAA 1587
Db 9182 TCTCTCAAAGTACCTCGAGAAAGTCAAGGCGCTTCGAGGAGAAATTCACCTCTGAAA 9241

Qy 1588 GAGAACTGAGCCACGCTCAATGACCTTGTCTGCGCAGCTTACCACTTTGGGCACTTCAGCTC 1647
Db 9242 GAGAACTGAGCTACGCTCAATGACCTTGTCTGCGCAGCTTCACTACGTTGGGCACTTCAGCTG 9301

Qy 1648 TCACCGTATTAACCTGAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCGAGGTG 1707
Db 9302 TCACCATATAACCTCAACACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCGAGGTG 9361

Qy 1708 GCCGTGAGGACCGAGTCAGGCACTGCTGAAGCCACAGGCACTTTGGTCAGGCACTCT 1767
Db 9362 GCCATTGAGGACCGCATCAGGCACTGCTGAAGCCACAGGCACTTTGGACCGCTCTC 9421

Qy 1768 CAGCACTTTCTTTCCAGCTCTGTCCAGGGTCTCTGGGAGAGAGCCATCTCGCCAAACAAA 1827
Db 9422 CAGCACTTTCTTTCCAGCTCTGTCTCGAGGGTCTCTGGGAGAGAGCCATCTCACCACAAA 9481

Qy 1828 GTGCCCTTACTATATCAACACGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACA 1887
Db 9482 GTGCCCTTACTATATCAACACGAGACCCAAACAACTTGTCTGGGACCATCCCAAAATGACA 9541

Qy 1888 GAGCTCTACCAAGTCTTTAGCTGACCTGAATTAATGTGATGATTTTCAAGCTTATAGACTGCC 1947
Db 9542 GAGCTCTACCAAGTCTTTAGCTGACCTGAATTAATGTGATGATTTTCAAGCTTATAGACTGCC 9601


```
Db 9467 GAGCTCTACCACTTTAGCTGACCTGATATATGTCAGGTTCTCGCGTATAGGACTGCC 9526
QY 1948 ATGAACCTCCGAAGACTGACAGAGCCCTTTGCTTGGATCTCTTGAGCTGTGACGTGCA 2007
Db 9527 ATGAAGCTCAGAAGGCTCCAGAAGGCCCTTTGCTTGGATCTCTTGAGCTGTGACGTGCA 9586
QY 2008 TGTGATGCTTGGACAGCAGACCACTCAAGCAAAATGACCAAGCCCATGGATATCTGCGAG 2067
Db 9587 TGTGATGCTTGGACAGCAGACCACTCAAGCAAAATGACCAAGCCCATGGATATCTGCGAG 9646
QY 2068 ATTATTAAATTTGTTGACCACTATTTATGACCCCTGAGAGCAAGAGCAACAATTTGGTC 2127
Db 9647 ATAATTAATCTTGTGACTACAAATTTATGATGCTGAGAGCAAGAGCAACAATTTGGTC 9706
QY 2128 AACGTCCTCTCTGGTGGATATGCTGAACTGGCTGCTGAATGTTTATGATACGGGA 2187
Db 9707 AATGTCCTCTCTGGTGGATATGCTGAACTGGCTGCTGAATGTTTATGATACGGGA 9766
QY 2188 CGAACAGGAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA 2247
Db 9767 CGAACAGGAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCTCTGTGTAAGCA 9826
QY 2248 CATTGGAAGACAGTACAGATACCTTTTCAGCAAGTGGCAAGTTCAACAGGATTTGT 2307
Db 9827 CACTTGGAAAGACAGTACAGATACCTTTTCAGCAAGTGGCAAGTTCAACTGGCTTTGT 9886
QY 2308 GACCAAGCAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCAGCAAGTGGGT 2367
Db 9887 GACCAAGCAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCAGCAAGTGGGT 9946
QY 2368 GAAGTTGCATCTTTGGGGGAGTAACTAGG 2401
Db 9947 GAAGTTGCATCTTTGGGGGAGTAACTAGG 9980

RESULT 15
AX538582
LOCUS 13815 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 2 from Patent WO0229056.
ACCESSION AX538582
VERSION AX538582.1 GI:25271088
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.O.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 2 11-APR-2002.
FEATURES
    Location/Qualifiers
        1..13815
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
BASE COUNT 4577 a 2719 c 3025 g 3494 t
ORIGIN
Query Match 62.8%; Score 1506.8; DB 6; Length 13815;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1607; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
QY 628 TGGCAACATCTGTAGATGACAGAGACCGGTGGTTCTTTACAGACACTCATAGA 687
Db 8207 TGGGAACATCTCATAAAGAGTAAGTGAAGAGAGCTGCTTTGGAAGAACTCATAGA 8266
QY 688 TTACTGCAACAGTTCCCTCGACCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAGCTGAA 747
Db 8267 TTACTGCAAGAGTTCCCTCGACCTGGACCTGGAAAAGTTTCTTGGCTGGCTTACAGAGCTGAA 8326
QY 748 ACAAAGTCCAAATGCTTACAGGATGCTACCGTGAAGAAAGGCTCTAGAGACTCCAG 807
```

```
Db 8327 ACAACTGCCAATGCTCTACAGAGCGCTTCCCTTAAGGAGAAAGCTCTTAGAAGACTCCAGG 8386
QY 808 GGGATTAAGAGAGCTGATGAACAATGGCAAGACCTCCCAAGGTGAATTTGAAGCTCACACA 867
Db 8387 GGGATCAGAGAGCTGATGAACAATGGCAAGACCTCCCAAGGTGAATTTGAAGCTCACACA 8446
QY 868 GATGTTTATCAACAACCTGGATGAACAAGCCAAAGAAATCTCTGAGATCCCTGGAGGTGCC 927
Db 8447 GATATCTATCAAACTTGTGATGAATAATGGCCAAAGAAATCTCTGAGATCCCTGGAGGTGCC 8506
QY 928 GATGATGCACTCTGTTTACAAAGAGCTTTGGATACATGAACCTTCAAGTGGAGTGAATTT 987
Db 8507 GATGAAGACACCCCTGTTTACAAAGAGCTTTGGATACATGAATTTCAAGTGGAGTGAATTT 8566
QY 988 CGGAAAAAGTCTCTCAACATTAGGTGCCATTTTGGAAAGCCAGTTCTGACCAAGTGGAGGCT 1047
Db 8567 CAGAAAAAGTCTCTCAACATTAGGTGCCATTTTGGAAAGCCAGTTCTGACCAAGTGGAGGCT 8626
QY 1048 CTGCACTTTCTCTGAGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGC 1107
Db 8627 TTGCATCTTTCTCTGAGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGC 8686
QY 1108 CGGCAAGCAGCTATTTGAGGCGACCTTTCCAGCAGCTTTCAGAAAGCAAGATGATGATAGC 1167
Db 8687 CGTCAAGCAGCTATTTGAGGCGACCTTTCCAGCAGCTTTCAGAAAGCAAGATGATGATAGC 8746
QY 1168 GCCTTCAAGAGGGAATTTGAAAGCTTAAAGAACTTAAAGAACTTAAAGAACTTAAAGAACT 1227
Db 8747 GCCTTCAAGAGGGAATTTGAAAGCTTAAAGAACTTAAAGAACTTAAAGAACTTAAAGAACT 8806
QY 1228 CGAATATTTCTGACAGAGAGCTTTTGGAAAGCAAGTGAAGAACTTACCAAGAGGCCAGA 1287
Db 8807 AGAATATTTCTGACAGAGAGCTTTTGGAAAGCAAGTGAAGAACTTACCAAGAGGCCAGA 8866
QY 1288 GAGCTGCTCTCTGAGAGAGAGCCCAAGATGTCAGTCCGCTTTTACGAAGCAGGCTGAG 1347
Db 8867 GAATGCTCTCTGAGAGAGAGCTTCAAGAACTTCAAGAACTTCAAGAACTTCAAGAACTGAA 8926
QY 1348 GAGTCAATACTGAGTGGGAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAATAAATA 1407
Db 8927 GAGTCAATACTGAGTGGGAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAATAAATA 8986
QY 1408 GATGAGACCTTTGAAAGACTTCCAGGAATTTCAAGAGGCCCAAGATGAGTGGACCTCAAG 1467
Db 8987 GATGAGACTTTGAAAGACTTCCAGGAATTTCAAGAGGCCCAAGATGAGTGGACCTCAAG 9046
QY 1468 CTGCGCAAGCTGAGGTGATCAAGAGGATCTGCGAGCCCGTGGGCGATCTCTCATTTGAC 1527
Db 9047 TTGCGCAAGCTGAGGTGATCAAGAGGATCTGCGAGCCCGTGGGCGATCTCTCATTTGAC 9106
QY 1528 TCTCTCCAGATCACTCCGAGAAAGTCAAGGCACTTTCGAGGAGAAATTTGGCTCTTGAAA 1587
Db 9107 TCTCTCCAGATCACTCCGAGAAAGTCAAGGCACTTTCGAGGAGAAATTTGGCTCTTGAAA 9166
QY 1588 GAGAACTGAGCCCACTCAATGACCTTGTGCGCAAGTTCACATTTTGGGCAATTCAGCTC 1647
Db 9167 GAGAACTGAGCCCACTCAATGACCTTGTGCGCAAGTTCACATTTTGGGCAATTCAGCTC 9226
QY 1648 TCACCTGATTAACCTCAGCACTCTGGAAGAGCTTCAACAGAGCTTTCGCAAGTGGAGCTG 1707
Db 9227 TCACCTGATTAACCTCAGCACTCTGGAAGAGCTTTCGCAAGTGGAGCTTTCGCAAGTGG 9286
QY 1708 GCGTTCGAGGACCGAGTTCAGGAGCTGCAATGAAGCCCAAGGAGCTTTGGTTCAGGATCT 1767
Db 9287 GCTGTGGAGGACCGTGTGACAGAGCTGCAATGAAGCCCAAGGAGCTTTGGTTCGATGCC 9346
QY 1768 CAGCACTTTCTTCCAGCTCTGTCAGGCTCCCTCGAGAGAGCCCATCTCGCCAAACAAA 1827
Db 9347 CAGCACTTTCTTCCAGCTCTGTCAGGCTCCCTCGAGAGAGCCCATCTCGCCAAACAAA 9406
QY 1828 GTGCGCTTACTATATCAACACAGAGACTTCAAAAGCTTGTGGGAGCCATCCCAAAATGACA 1887
```

Db	9407	GTGCCCTACTATATCAACCAGAGACCCAAACCACTTGTGGAGCACCCCAAAATGACA	9466
Qy	1888	GAGCTCTACAGTCTTTAGCTGACCTGATTAATAGTCAGATTCTCAGCTTATAGGCTGCC	1947
Db	9467	GAGCTCTACAGTCTTTAGCTGACCTGATTAATAGTCAGATTCTCAGCTTATAGGCTGCC	9526
Qy	1948	ATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCA	2007
Db	9527	ATGAAGCTCAGAAGGCTCCAGAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGCA	9586
Qy	2008	TGTGATGCTTGGACGACGACCAACCTCAAGCAAAATGACAGAGCCCATGATATCCTGAG	2067
Db	9587	TGTGATGCTTGGACGACGACCAACCTCAAGCAAAATGACAGAGCCCATGATATCCTGAG	9646
Qy	2068	ATTATTAAATGTTTGACCACTATTATGACCGCTGGAGCAAGAGCACAACTTTGGTC	2127
Db	9647	ATAATTAACTGTTTGACTACAATTTATGATCGTCTGGAGCAAGAGCACAACTTTGGTC	9706
Qy	2128	AAGTCCCTCTCTGCGTGGATATGCTGTAAGTGGCTGCTGTAATGTTTATGATACGGGA	2187
Db	9707	AATGTCCCTCTCTGCTGGATATGCTGTAAGTGGCTGCTGTAATGTTTATGATACGGGA	9766
Qy	2188	CGAACAGGAGGATCCGTGCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA	2247
Db	9767	CGAACAGGAGGATCCGTGCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAGCA	9826
Qy	2248	CATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT	2307
Db	9827	CACCTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT	9886
Qy	2308	GACCAGCGAGGCTGGGCTCCCTCTGTCATGTTCTATCCAAATCCAAAGACAGTTGGGT	2367
Db	9887	GACCAGCGAGGCTGGGCTCCCTCTGTCATGTTCTATCCAAATCCAAAGACAGTTGGGT	9946
Qy	2368	GAAGTTGCATCCTTTGGGGCAGTAACATTGAGC	2401
Db	9947	GAAGTTGCATCCTTTGGGGCAGTAACATTGAGC	9980

Search completed: February 1, 2004, 23:51:05
 Job time : 8561.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:04:35 ; Search time 3752.09 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000

Perfect score: 2001

Sequence: 1 ggcagtcattgatgagag.....tcaaccacagactcaaca 2001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estm:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hrc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_ptg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	678.4	33.9	3870	11	BC036103
2	591.8	29.6	834	12	BI729851
3	547.4	27.4	3056	11	AK044536
4	491.4	24.6	750	12	BI730168

	5	486.6	24.3	579	9	AL121550
	6	471.6	23.6	728	14	CB228986
	7	424.2	21.2	663	12	BM488464
	8	414.6	20.7	1490	11	BC009242
c	9	409	20.4	595	14	CB177816
	10	389	19.4	644	13	EU313510
c	11	324.6	16.2	402	14	CB547284
	12	309.8	15.5	3753	11	AK081426
	13	265.4	13.3	3051	11	BC036095
	14	261.2	13.1	2334	11	BC011062
	15	257.8	12.9	1541	11	AK034383
	16	252.4	12.6	772	12	BI250598
	17	233	11.6	835	12	BI553820
	18	230.8	11.5	784	10	EG212445
	19	230.2	11.5	763	14	CB518960
	20	229.2	11.5	778	14	CB524596
	21	226	11.3	250	10	BF963618
	22	216.6	10.8	681	13	BU301653
	23	199.8	10.0	696	12	BJ075057
	24	199.2	10.0	697	13	EU456556
c	25	182.4	9.1	645	9	AW467977
	26	179.6	9.0	466	14	CD549993
	27	172.2	8.6	642	9	AL855376
	28	167.6	8.4	502	9	AL602076
c	29	167.2	8.4	646	12	BI289102
	30	165.4	8.3	851	13	BU201022
	31	164.6	8.2	600	12	BI988528
	32	158	7.9	405	14	CB811308
	33	153.4	7.7	481	10	BE557463
	34	153.2	7.7	885	11	AK020881
c	35	150.6	7.5	515	28	AZ780914
	36	144.6	7.2	675	14	BY706879
	37	144.2	7.2	392	10	AW948794
	38	142.2	7.1	426	13	EQ319056
	39	141.6	7.1	794	28	BZ097143
c	40	141	7.0	761	28	BZ129893
	41	140.6	7.0	649	13	BQ209933
c	42	140	7.0	515	10	BB651473
	43	137.8	6.9	479	13	BQ304046
c	44	134	6.7	828	28	BZ126085
	45	133.2	6.7	521	28	AZ226551

ALIGNMENTS

```
RESULT 1
BC036103
LOCUS BC036103 3870 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:5274415, mRNA.
ACCESSION BC036103
VERSION BC036103.1 GI:23271310
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3870)
Strausberg, R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
```

NIH-MGC Project URL: http://mgc.nci.nih.gov

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@axil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 48 Row: f Column: 7
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: retained intron.

FEATURES

source

1. 3870
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5274415"
 /tissue_type="Brain, hippocampus"
 /clone_lib="NIH MGC_95"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

BASE COUNT 1321 a 677 c 763 g 1103 t

ORIGIN

Query Match 33.9%; Score 678.4; DB 11; Length 3870;
 Best Local Similarity 97.7%; Pred. No. 1.9e-126;
 Matches 688; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCAGTTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
 Db 1143 GGCAGTTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1202

QY 61 GTATTATCGTGCTTTCTCTCGAGGACACATTCGAAGCAAGAGAGAGATTCTTAAT 120
 Db 1203 GTATTATCGTGCTTTCTCTCGAGGACACATTCGAAGCAAGAGAGAGATTCTTAAT 1262

QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGTACATGATGATTGACA 180
 Db 1263 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGTACATGATGATTGACA 1322

QY 181 GCCATTCAGGGCGGGTGGTGAATTTCTACAAATTTGGGAAGTAAGCTGATTGGAACAGGA 240
 Db 1323 GCCATTCAGGGCGGGTGGTGAATTTCTACAAATTTGGGAAGTAAGCTGATTGGAACAGGA 1382

QY 241 AAATTATCAGAGATGAGAACTGAAGTACAAAGACAGATGAATCTCTTAATTCAGA 300
 Db 1383 AAATTATCAGAGATGAGAACTGAAGTACAAAGACAGATGAATCTCTTAATTCAGA 1442

QY 301 TGGCAATGCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
 Db 1443 TGGCAATGCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 1502

QY 361 GATCTCCAGATCAGAAACTGAAGAGTTGAATGATCTGCTTAACAAAAACAGAGAAAGA 420
 Db 1503 GATCTCCAGATCAGAAACTGAAGAGTTGAATGATCTGCTTAACAAAAACAGAGAAAGA 1562

QY 421 ACAAGGAATGAGGAAGGCTCTTGACCTGATCTTTGAGAGCTTAAAGCGCAAGTA 480
 Db 1563 ACAAGGAATGAGGAAGGCTCTTGACCTGATCTTTGAGAGCTTAAAGCGCAAGTA 1622

QY 481 CAACAACATAGGTGCTTCAAGAGATCTAGAACAGAACTAGTCAAGGTCAATTTCTCTC 540
 Db 1623 CAACAACATAGGTGCTTCAAGAGATCTAGAACAGAACTAGTCAAGGTCAATTTCTCTC 1682

QY 541 ACTCAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
 Db 1683 ACTCAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1742

QY 601 GAACAACATAGGTGCTTCAAGAGATCTAGAACAGAACTAGTCAAGGTCAATTTCTCTC 660
 Db 1743 GAACAACATAGGTGCTTCAAGAGATCTAGAACAGAACTAGTCAAGGTCAATTTCTCTC 1802

QY 661 TGGGTTCTTTTACAAGACAGCCTGACCTAGCTAGCTCTCTGGAGCTGAC 704
 Db 1803 TGGGTTCTTTTACAAGACATCTCTCTCAATGGCAACGCTCTTAC 1846

RESULT 2

BI729851

LOCUS

DEFINITION 603349511F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5', mRNA sequence.

ACCESSION

BI729851

VERSION

BI729851.1

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 834)

NIH-MGC <http://mgi.nci.nih.gov/>

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabs@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11908 row: e column: 03

High quality sequence stop: 796.

Location/Qualifiers

1. 834

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:5357162"

/tissue_type="retina"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 94"

/notes="Organ: eye; Vector: pCMV-Sport6; Site: 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 3.3 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 275 a 173 c 199 g 187 t

ORIGIN

Query Match

Best Local Similarity

Matches 665; Conservative 0; Mismatches 67; Indels 3; Gaps 3;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTCTTCTGCTG 872

Db 75 GGAAGAATCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAAGTTCTTCTGCTG 134

QY 873 GCTTACAGAGCTGAAACCACTGCCAATGTCTTACAGGATGCTACCCGTAAGGAAGGCT 932

Db 135 GATTACGGAAGCAGAAACCACTGCCAATGTCTTACAGGAGCTTCCCGTAAGGAGAGCT 194

QY 933 CCTAGAAGACTCAAGGGGATAAAGAGCTGTATGAACCAATGGCAAGACCTCCAGGTGA 992

Db 195 CCTAGAAGACTCC-AGGAGTCAAGAGCTGTATGAACCAATGGCAAGATCTCCAAGGAGA 253

QY 993 AATTGAGCTCACACAGATCTTTATCACACCTGGATGAAACAGCCAAAAATCTGTAG 1052

Db 254 AATTGAAATCTCACACAGATCTTTATCACAACTTTGATGAAATGGCCAAAAATCTGTAG 313

QY 1053 ATCCCTGGGAAGGTTCCGATGATGAGTCTCTGTTTACAAAGACGTTTGGATTAACATGACTT 1112

Db 314 ATCCCTGGGAAGGTTCCGATGATGAGTCAAGCACCCCTGTTTACAAAGACGTTTGGATTAACATGACTT 373


```
misc_feature 1. 3056
/note="dystrophin, muscular dystrophy (MCD|MGT:94909,
GB|NW_007868, evidence: BLASTN, 100%, match=99%)
BASE COUNT 1072 a 567 c 607 g 810 t
ORIGIN
Query Match 27.4%; Score 547.4; DB 11; Length 3056;
Best Local Similarity 88.0%; Pred. No. 4.9e-100; Indels 0; Gaps 0;
Matches 596; Conservative 0; Mismatches 81;
QY 1 GCGAGTTCATTGATGGAGAGTGAAGTAAACCTGACCGCTTATCAACAGCTTTAGAGAA 60
DB 1044 CACAGTTCATTGATGGAGAGTGAAGTAAACCTGATAGTTACCAAACTGCTTTAGAGAA 1103
QY 61 GTATTATCGTGGCTCTTCTTCGTGGAGACACATTCGAAGCAGAGAGAGATTTCTAAT 120
DB 1104 GTACTTCATGGCTCTTCTTCGCGAGATACATTCGAGCACAGAGAGATTTCAAT 1163
QY 121 GATGTGGAAGTGGTGAAGACCACTTTTATCTACTCATGAGGGTACATGATGGATTTGACA 180
DB 1164 GATGTGGAAGTGGTGAAGACCACTTTTATCTCATGAGGGATTCATGATGGATCTGACA 1223
QY 181 GCCATCAGGCGCGGTTGGTAAATATCTTCAATTTGGAGTAACTGATGGAGAGGA 240
DB 1224 TCTCATCAAGGACTGTGTGGTAAATTTCTACAGTTAGGATCACTAGTTGGAAAGGG 1283
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA 300
DB 1284 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAAATGAATCTCTAAATTTCAAGA 1343
QY 301 TGGGAATCCCTCAGGTTAGTACATGAGGAAACAAAGCAATTTACATAGATTTTAATG 360
DB 1344 TGGGAATCTCTCAGGTTAGTACATGAGGAAACAAAGCAATTTACAAAGTTCTAATG 1403
QY 361 GATCTCCAGATCAGAACTCAAGAGTTGAATGACTGGCTTAACAAACAAAGAGAAAGA 420
DB 1404 GATCTCCAGATCAGAACTCAAGAGTTGAATGACTGGCTTAACAAACAAAGAGAGAG 1463
QY 421 ACAAGAAATAGGAGAGAGAGCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 480
DB 1464 ACTAAGAAATAGGAGAGAGAGCTTTGGACCTGATCTTGAAGATCTTAAATGCGAAGTA 1523
QY 481 CAACAACTAAGTCTCTTCAAGAGATCTAGAAACAAAGTCAAGGTCATTTCTCTC 540
DB 1524 CAACAACTAAGTCTCTTCAAGAGATCTAGAAACAAAGTCAAGGTCATTTCTCTC 1583
QY 541 ACTCATGTTGGTGGTGGTATGATGAATCTAGTGGAGATCAGCAACTGCTGTTTGGAA 600
DB 1584 ACTCATGTTGGTGGTGGTATGATGAATCTAGTGGAGATCAGCAACTGCTGTTTGGAA 1643
QY 601 GAACAACTAAGTGGAGATGATGGGCAACATCTCTAGATGGACAGAGAGCCG 660
DB 1644 GAACAACTAAGTGGAGATGATGGGCAACATCTCTAGATGGACAGAGAGCCG 1703
QY 661 TGGGTTCTTTTACAAGA 677
DB 1704 TGGATTGTTTACAAGA 1720

RESULT 4
BI730168
LOCUS 603349711F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357187 5',
DEFINITION mRNA sequence.
ACCESSION BI730168
VERSION BI730168.1 GI:15707181
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 750)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
```

```
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11908 row: f column: 04
High quality sequence stop: 747.
FEATURES
Location/Qualifiers
1..750
source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357187"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
BASE COUNT 240 a 161 c 178 g 170 t 1 others
ORIGIN
Query Match 24.6%; Score 491.4; DB 12; Length 750;
Best Local Similarity 88.4%; Pred. No. 8.3e-89; Indels 8; Gaps 6;
Matches 602; Conservative 0; Mismatches 71;
QY 804 GATGTTGGAGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAGTT 863
DB 69 GCTGTTGGACAGAACTCATAGATTACTGCAGCAGTTCCTCTGGACCTGGAGAGTT 128
QY 864 TCTTGCTGG-CTTACAGAGCTGAAAC--AACTGCCAATGTCTTACAGGATGCTACCCG 920
DB 129 TCTTCTCTGGATTACGGAAGCAGAAACAGAGCTGCCAATGTCTTACAGGACGCTTCCG 188
QY 921 -TAAGAAAGGCTCTTAGAAGACTCCAAAGGG--AGTAAAGAGCTGATGAACAAATGGCA 977
DB 189 CTAAGGAGAGCTCTTAGAAGCTCCAGGGGAGTCAAGAGCTGATGAACCAATGGCA 248
QY 978 AGACCTCCAAGTGAAA--TTGAGCTCAACAGATGTTTATCAACCTGATGAACAA 1036
DB 249 AGATCTCCAAGGAGAAAGTTGAAACTCAACAGATATCTATCAAACTTTGATGAAGATG 308
QY 1037 GCCAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTTCAAAAGACGTT 1096
DB 309 GCGAAAAATCTGAGATCCCTGGAAGGTTGGATGAAGCACCCTGTTTCAAGAGCGTT 368
QY 1097 TGGATAACAGAACTTCAAGTGGAGTGAACCTTGGAAAAAGTCTTCAACATAGGTCCC 1156
DB 369 TGGATAACAGAACTTCAAGTGGAGTGAACCTTCAAAAAAGTCTTCAACATAGGTCCC 428
QY 1157 ATTTGGAGCCAGTTCTGACAGTGGAGGCTGTCACCTTTCTCTGAGAACTTTCTGG 1216
DB 429 ATTTGGAGCCAGTTCTGACAGTGGAGGCTTGCATCTTTCTCTGAGAACTTTCTGG 488
QY 1217 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGCGACCTTATTGGAGCGCACTTTC 1276
DB 489 TTTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGCGACCTTATTGGAGCGCACTTTC 548
QY 1277 CAGCAGTTCAGAGCA--GAACGATGATAGGCGCTTCAAGAGGGAATTTGAAACTAAA 1335
DB 549 CAGCAGTTCAGAGCAGCAATGATATACATAGGCGCTTCAAGAGGGAATTTGAAACTAAA 608
QY 1336 GAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTTCTGACAGAGCAGCTTTG 1395
DB 609 GAACCTGTAAATCATGAGTACTCTTGAGACTGTGAGAATATTTTCTGACAGAGCAGCTTTG 668
```

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov

Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LLAM11908 row: f column: 04
High quality sequence stop: 747.

FEATURES

Location/Qualifiers
1..750
source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357187"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."

BASE COUNT 240 a 161 c 178 g 170 t 1 others

Query Match 24.6%; Score 491.4; DB 12; Length 750;

Best Local Similarity 88.4%; Pred. No. 8.3e-89; Indels 8; Gaps 6;
Matches 602; Conservative 0; Mismatches 71;

QY 804 GATGTTGGAGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAGTT 863

DB 69 GCTGTTGGACAGAACTCATAGATTACTGCAGCAGTTCCTCTGGACCTGGAGAGTT 128

QY 864 TCTTGCTGG-CTTACAGAGCTGAAAC--AACTGCCAATGTCTTACAGGATGCTACCCG 920

DB 129 TCTTCTCTGGATTACGGAAGCAGAAACAGAGCTGCCAATGTCTTACAGGACGCTTCCG 188

QY 921 -TAAGAAAGGCTCTTAGAAGACTCCAAAGGG--AGTAAAGAGCTGATGAACAAATGGCA 977

DB 189 CTAAGGAGAGCTCTTAGAAGCTCCAGGGGAGTCAAGAGCTGATGAACCAATGGCA 248

QY 978 AGACCTCCAAGTGAAA--TTGAGCTCAACAGATGTTTATCAACCTGATGAACAA 1036

DB 249 AGATCTCCAAGGAGAAAGTTGAAACTCAACAGATATCTATCAAACTTTGATGAAGATG 308

QY 1037 GCCAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGCAGTCTCTTCAAAAGACGTT 1096

DB 309 GCGAAAAATCTGAGATCCCTGGAAGGTTGGATGAAGCACCCTGTTTCAAGAGCGTT 368

QY 1097 TGGATAACAGAACTTCAAGTGGAGTGAACCTTGGAAAAAGTCTTCAACATAGGTCCC 1156

DB 369 TGGATAACAGAACTTCAAGTGGAGTGAACCTTCAAAAAAGTCTTCAACATAGGTCCC 428

QY 1157 ATTTGGAGCCAGTTCTGACAGTGGAGGCTGTCACCTTTCTCTGAGAACTTTCTGG 1216

DB 429 ATTTGGAGCCAGTTCTGACAGTGGAGGCTTGCATCTTTCTCTGAGAACTTTCTGG 488

QY 1217 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGCGACCTTATTGGAGCGCACTTTC 1276

DB 489 TTTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGCGACCTTATTGGAGCGCACTTTC 548

QY 1277 CAGCAGTTCAGAGCA--GAACGATGATAGGCGCTTCAAGAGGGAATTTGAAACTAAA 1335

DB 549 CAGCAGTTCAGAGCAGCAATGATATACATAGGCGCTTCAAGAGGGAATTTGAAACTAAA 608

QY 1336 GAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTTCTGACAGAGCAGCTTTG 1395

DB 609 GAACCTGTAAATCATGAGTACTCTTGAGACTGTGAGAATATTTTCTGACAGAGCAGCTTTG 668

```

QY 1396 GAAGGACTAGAGAACTTACACAGGAGCCAGAGAGCTGCCTCTGAGGAGAGCCAG 1455
    |||
Db 669 GAAGGACTAGAGAACTTACACAGGAGCCAGAGAGCTGCCTCTGAGGAGAGCTAG 728
    |||
2Y 1456 AATGTCACCTGGCTTCTACGA 1476
    |||
3b 729 AATGTCACCTGGCTTCTACGA 749
    |||

RESULT 5
LOCUS AL121550
DEFINITION DKFP762L078_r1.762 (synonym: hm12) Homo sapiens cDNA clone
ACCESSION AL121550
VERSION AL121550.1 GI:5927551
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 579)
AUTHORS Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished
COMMENT Contact: Ottenwaelder B
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ) | Email s.wiemann@dkfz-heidelberg.de;
sequenced by MedGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFP762L078) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
    source
        1..579
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFP762L078"
            /tissue_type="melanoma (MeWo cell line)"
            /dev_stage="adult"
            /lab_host="DH10B"
            /notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 180 a 128 c 144 g 127 t
ORIGIN
Query Match 24.3%; Score 486.6; DB 9; Length 579;
Best Local Similarity 97.2%; Pred. No. 7.5e-88;
Matches 495; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

2Y 977 AGACCTCCAGGTGAATAGCTCACACAGATGTTTATCACACCTGGATGAACA 1036
    |||
3b 71 AGACCTCCAGGTGAATAGCTCACACAGATGTTTATCACACCTGGATGAACA 130
    |||
2Y 1037 GCCAAAAATCTTGAGATCCCTGGAAAGGTTCCGATGATCGAGTCTGTATACAAAGACGTT 1096
    |||
3b 131 GCCAAAAATCTTGAGATCCCTGGAAAGGTTCCGATGATCGAGTCTGTATACAAAGACGTT 190
    |||
2Y 1097 TGGATACATGAATCAATGAGTGAATCTTCGGAAAAAGTCTCTCAACATTAGGTCCC 1156
    |||
3b 191 TGGATACATGAATCAATGAGTGAATCTTCGGAAAAAGTCTCTCAACATTAGGTCCC 250
    |||
2Y 1157 ATTTGGAAGCCAGTCTTGACACCGTGAACGCTGACACCTTTCTCTGGAGGAACCTTCGG 1216
    |||
3b 251 ATTTGGAAGCCAGTCTTGACACCGTGAACGCTGACACCTTTCTCTGGAGGAACCTTCGG 310
    |||

```

```

QY 1217 TGTGGCTACAGCTGAAGATGATGAATTAAGCCGAGGACCACTATTGGAGGCACTTTC 1276
    |||
Db 311 TGTGGCTACAGCTGAAGATGATGAATTAAGCCGAGGACCACTATTGGAGGCACTTTC 370
    |||
QY 1277 CAGCAGTTTCAGAGCAGACGATGTACATAGGGCTTTCAAGAGGGAAATTTGAAACTAAAG 1336
    |||
Db 371 CAGCAGTTTCAGAGCAGACGATGTACATAGGGCTTTCAAGAGGGAAATTTGAAACTAAAG 430
    |||
QY 1337 AACCTGTAATCATGAGTACTCTTGAGACTCTAGCAATATTTCTGACAGAGCAGCTTTGG 1396
    |||
Db 431 AACCTGTAATCATGAGTACTCTTGAGACTCTAGCAATATTTCTGACAGAGCAGCTTTGG 490
    |||
QY 1397 AAGGACTAGAGAACTCTTACAGAGCCAGAGAGTGCCTCTCTGAGAGAGAGCCAGAG 1456
    |||
Db 491 AAGGCTTAGAGAACTCTTACAGAGCCAGAGAGTGCCTCTCTGAGAGAGAGCCAGAG 550
    |||
QY 1457 ATGTCACCTCGGCTCTTACGAAACGAGCT 1485
    |||
Db 551 ATGTCACCTCGCTTCTACGAAACGAGCT 579
    |||

RESULT 6
LOCUS CB228986
DEFINITION AGENCOURT_11499247 NICHDRh_Ov1 Macaca mulatta cDNA clone
IMAGE:6884820 5', mRNA sequence.
ACCESSION CB228986
VERSION CB228986.1 GI:28280564
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
REFERENCE 1 (bases 1 to 728)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: sgapbs@mail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
cDNA Library Preparation: CLONTECH
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM3135 row: e column: 11
High quality sequence stop: 583.

FEATURES
    source
        1..728
            /organism="Macaca mulatta"
            /mol_type="mRNA"
            /db_xref="taxon:9544"
            /clone="IMAGE:6884820"
            /tissue_type="Ovary"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NICHDRh_Ov1"
            /notes="Organ: ovary; Vector: pDNR-LiB; Site 1: Sfi I;
            Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.0-4.0 kb. Tissue pooled from
            pre-pubertal, post pubertal an menopausal monkeys.
            Constructed by Clontech. Note: this is a NICHDRh Library."

BASE COUNT 256 a 127 c 162 g 182 t
ORIGIN
Query Match 23.6%; Score 471.6; DB 14; Length 728;
Best Local Similarity 95.3%; Pred. No. 8.1e-85;
Matches 486; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 195 GGTGGTATATTTCTACATTTGGAGTAGCTGATGGAACAGGAAATTTATCAGAGA 254
    |||

```

Db 1 GGTGGTAAATCTTCAATTCGGAGTCAGCTGATTGGAACAGGGAATATCAGAGA 60
 QY 255 TGAAGAACTGAAGTACAGAGCAGATGATCTCTTAATTAACAGATGGGAATGCTCAG 314
 Db 61 TGAAGAACTGAAGTACAGAGCAGATGATCTCTTAATTAACAGATGGGAATGCTCAG 120
 QY 315 GGTAGCTAGCTGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCGAATCA 374
 Db 121 GGTAGCTAGCTGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCGAATCA 180
 QY 375 GAACTGAAGAGTGAATGATGCTGCTAACAAAACAGAGAAGAAACAAGAAAATGGA 434
 Db 181 GAACTGAAGAGTGAATGATGCTGCTAACAAAACAGAGAAGAAACAAGAAAATGGA 240
 QY 435 GGAAGAGCTCTTGACCTGATCTTGAGACCTAAACGCCAAGTACAAACATAGGT 494
 Db 241 GGAAGAGCTCTTGACCTGATCTTGAGACCTAAACGCCAAGTACAAACATAGGT 300
 QY 495 GCTTCAAGAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCACTCATAGTGGT 554
 Db 301 GCTTCAAGAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCACTCATAGTGGT 360
 QY 555 GGTAGTGTATGATCTAGTGGAGATCAGCACTGCTGCTTTGGAAGAACATTAAGT 614
 Db 361 GGTAGTGTATGATCTAGTGGAGATCAGCACTGCTGCTTTGGAAGAACATTAAGT 420
 QY 615 ATTGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTACA 674
 Db 421 ATTGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTACA 480
 QY 675 AGACAGCTGACCTAGCTCTGGAGCTGAC 704
 Db 481 AGACATCTTCTCAATGGCAAGCTCTTAC 510

RESULT 7
 BM488464
 LOCUS
 DEFINITION
 pgm2n.pk007.117 Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA clone pgm2n.pk007.117 5' similar to sp|P11533|DMD CHICK Dystrophin p1r|S02041 dystrophin, muscle - chicken emb|CAA31746.1| (X13369) dystrophin (AA 1 - 3660) [Gallus gallus], mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Gallus gallus (chicken)
 EST.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 Reference 1 (bases 1 to 663)
 Cogburn, L.A. and Monsonego-Ornan, E.
 ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IRAFs Animal Genome Project

JOURNAL
 COMMENT
 Contact: Larry A. Cogburn
 University of Delaware
 Townsend Hall, Newark, DE 19717, USA
 Tel: 302-831-1335
 Fax: 302-831-2822
 Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES
 source

1. .663
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Commercial broiler and Ottawa Res. Centre Strains 90 & 21"
 /db_xref="taxon:9031"
 /closes="pgm2n.pk007.117"
 /sex="Male and female"
 /tissue_type="Breast muscle, leg muscle and epiphyseal

RESULT 8
 BC009242
 LOCUS
 DEFINITION

Homo sapiens, dystrophin (muscular dystrophy, Duchenne and Becker types), includes DXS142, DXS164, DXS206, DXS230, DXS268, DXS269, DXS270, DXS272, clone IMAGE:3029414, mRNA.

growth plate"
 /dev_stage="Breast, leg:Embryo(d19);post-hatch(1d,1.3,5,7,9,11 weeks);growth plate(1d,7d,14d post-hatch)"
 /lab_host="E. coli EMDH10B"
 /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm2n)"
 /note="Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each tissue (embryonic muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth plate 33.3% of the final RNA pool). Single pass sequencing from 5'-end"

BASE COUNT 209 a 132 c 172 g 144 t 6 others
 ORIGIN

Query Match 21.2%; Score 424.2; DB 12; Length 663;
 Best Local Similarity 78.1%; Pred. No. 2.9e-75;
 Matches 518; Conservative 0; Mismatches 144; Indels 1; Gaps 1;

QY 960 GCTGATGAACAATGGCAAGACCTCCAGGTGAATTTGAAGCTCACAGATGTTTATCA 1019
 Db 1 GCTCATGAAGCAGTGGCAGGATCTACAGGCAGAAATTTGATGCACATCTGACATCT-TNN 59
 QY 1020 CAACCTGGATGAAGCAAGCAAAATCTCGATCCCTGGAGGTTCCGATGATGCAGCT 1079
 Db 60 CAACCTGGATGAAGCAAGCAAAATCTCGATCCCTGGAGGTTCCGATGATGCAGCT 119
 QY 1080 CCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTCGAGTGAACCTTCGAAAAAGTC 1139
 Db 120 CCTGTTGCAGACGCTCTGGATAACATGAACCTTCAAGTCGAGTGAACCTTTCAGGAAGATC 179
 QY 1140 TCTCAACATTAGTCTCCCATTTGGAAGCCAGTTCTGACAGTGGAGCGTCTGCACCTTTC 1199
 Db 180 TCTAAACATTAGATCTCTTTGGAAGCCAGCAGACCCAGTGGAGCGTTTACATCTCTC 239
 QY 1200 TCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCGCGGAGCACC 1259
 Db 240 TCTTCAGGAACCTTTCGGCATGGCTGCAATTTGAAGGAGGATGAATTAACAGCAAGCACC 299
 QY 1260 TATTGGAGGCGACTTTCAGCAGCTTCCAGAGCAGAACGATGTACATAGGCGCTTCAAGAG 1319
 Db 300 CAITGGTGGAGATATCCCACTGTGCAAGCAGAAATGATTTTCATAGGACTTTCAGAGAG 359
 QY 1320 GGAATTCAAAACATAAGAACCTGTATCATGAGTACTCTTTGAGACTGTACGAATATTTCT 1379
 Db 360 GGAGCTGAAAACAAAGAGACCTGTTATCATGAATGCACCTTGAGACTGTGCGACTCTTCT 419
 QY 1380 GACAGAGCAGCTTTTGAAGAGCTTAGAGAACTCTTACCAGAGCCCGCAGAGAGTGCCTCC 1439
 Db 420 GGCAGATCAACACAGTAGAGGACTTGGAAAAGGTCTATCCAGAACCAAGAGACCTATCACC 479
 QY 1440 TGAGGAGAGAGCCCGAATGCTCACTCGGCTTCTACGAAGCAGCGCTGAGAGGTCAATAC 1499
 Db 480 TGAGGAGAGGCGCCAGAAATGTCACTAAAGTTCTCCGAAGCAGAGAGATGATGTCAAGAC 539
 QY 1500 TGAGTGGAAAAAATTGAACCTGCACCTCGCTGACTGCGCAGAGAAAAATAGATGAGACCT 1559
 Db 540 TGAGTGGGAATAGCTAAATCTACGTTCTGCTGATTGGCAAAAGAGATAGATGCTCT 599
 QY 1560 TGAAGAAGCTCCAGAACTTCAAGAGCCCGAGATGAGCTGGAGCTCAAGCTGCGCCAGC 1619
 Db 600 TGAAGAAGCTCAGGGTCTTCAGGAGGCAATGGATGAACCTNNNCTGAAAGTGGCGCCAGGC 659
 QY 1620 TGA 1622
 Db 560 TGA 662

ACCESSION BC009242
 VERSION BC009242.1 GI:14714379
 KEYWORDS HTC
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 TITLE Direct Submission
 AUTHORS Strausberg, R.
 JOURNAL Submitted (06-JUN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadansystemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettenan, Anuradha
 Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 10 Row: j Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 5032284
 This clone has the following problem: retained intron.

FEATURES
 source
 1..1490
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3029414"
 /tissue_type="Muscle, rhabdomyosarcoma"
 /clone_lib="NIH_MGC_17"
 /lab_host="DH10B-R"
 /notes="Vector: pOTB7"
 BASE COUNT 505 a 299 c 328 g 358 t
 ORIGIN

Query Match 20.7%; Score 414.6; DB 11; Length 1490;
 Best Local Similarity 99.0%; Pred. No. 2.8e-73;
 Matches 417; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 GGCAGTTTCATTGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
 Db 1069 GGCAGTTTCATTGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1128
 Qy 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGACGACACAGAGATTTCTAAT 120
 Db 1129 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGACGACACAGAGATTTCTAAT 1188
 Qy 121 GATGTGGAAGTGTGAAGACACAGTTTCATCATGAGGGTACATGATGATTTCACA 180
 Db 1189 GATGTGGAAGTGTGAAGACACAGTTTCATCATGAGGGTACATGATGATTTCACA 1248
 Qy 181 GCCCATCAGGCCGGGTGGTAATATTCTACAAATGGGAAGTGAAGTGAAGACAGGA 240
 Db 1249 GCCCATCAGGCCGGGTGGTAATATTCTACAAATGGGAAGTGAAGTGAAGACAGGA 1308
 Qy 241 AATATTACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 300
 Db 1309 AATATTACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1368
 Qy 301 TGGGAATCCCTCAGGGTAGCTAGCATGGGAAAAACAAGCAATTTACATAGATTTTAATG 360
 Db 1369 TGGGAATCCCTCAGGGTAGCTAGCATGGGAAAAACAAGCAATTTACATAGATTTTAATG 1428

Qy 361 GATCTCCAGATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGA 420
 Db 1429 GATCTCCAGATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGA 1488
 Qy 421 A 421
 Db 1489 A 1489

RESULT 9
 CBI77816/C
 LOCUS CBI77816 595 bp mRNA linear EST 31-JAN-2003
 DEFINITION is21c01.xl HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3
 similar to SW:DMW_HUMAN P11532 DYSTROPHIN. [1]; mRNA sequence.
 ACCESSION CBI77816
 VERSION CBI77816.1 GI:28186206
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 595)
 AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lenishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Page, D., Wyllie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas,
 M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
 Jackson, Y., and Bowers, Y.
 Other ESTs: is21c01.y1
 Unpublished
 Endocrine Pancreas Consortium
 TITLE CBI77816
 JOURNAL
 COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@bioh.harvard.edu
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoe@im.wustl.edu)
 Seg primer: -40RP from Gibco
 High quality sequence stop: 448.
 Location/Qualifiers
 1..595
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6553129"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85 islet"
 /notes="Organ: Pancreas; Vector: pBluescript SK(-); Site: 1;
 NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
 Size selected on agarose gel. Average insert size ~1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoe@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."

BASE COUNT 146 a 136 c 116 g 197 t
 ORIGIN
 Query Match 20.4%; Score 409; DB 14; Length 595;
 Best Local Similarity 100.0%; Pred. No. 3.4e-72;
 Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GGCAGTTTCATTGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
 Db 409 GGCAGTTTCATTGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 350

QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACATTTGCAAGCACAAGGAGAGATTCTTAAT 120
 Db 349 GATATTATCGTGGCTTCTTCTGCTGAGGACATTTGCAAGCACAAGGAGAGATTCTTAAT 290
 QY 121 GATGTGAAGTGGTGAAGACCAAGTTCATCTCATGAGGGGTACATGATGATTGACA 180
 Db 289 GATGTGAAGTGGTGAAGACCAAGTTCATCTCATGAGGGGTACATGATGATTGACA 230
 QY 181 GCCATCAGCGCGGGTGGTGAATTTCTACAAATTTGGGAAGTAAAGCTGATTGGAACAGGA 240
 Db 229 GCCATCAGCGCGGGTGGTGAATTTCTACAAATTTGGGAAGTAAAGCTGATTGGAACAGGA 170
 QY 241 AATATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGA 300
 Db 169 AATATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGA 110
 QY 301 TGGGAATGCTCAGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAATG 360
 Db 109 TGGGAATGCTCAGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAATG 50
 QY 361 GATCTCCAGAAATCAGAAATCGAAAGAGTTGAATGACTGGCTAACAAAAA 409
 Db 49 GATCTCCAGAAATCAGAAATCGAAAGAGTTGAATGACTGGCTAACAAAAA 1

RESULT 10
 BUI313510
 LOCUS 644 bp mRNA linear EST 28-NOV-2002
 DEFINITION 603540290F1 CSEQCHN61 Gallus gallus cDNA clone CHEST508C24 5', mRNA sequence.
 ACCESSION BUI313510
 VERSION BUI313510.1 GI:25921511
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 1 (bases 1 to 644)
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 2235534
 12445392
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1. .644
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Compton line 151"
 /db_xref="taxon:9031"
 /clone="CHEST508C24"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH103"
 /clone_lib="CSEQCHN61"
 /note="Organ: heart; Vector: pBluescript II KS(+); Site_1: EcRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended ligated to NotI adapters, digested with EcRI, size-selected, and cloned into the NotI and EcRI

FEATURES

source
 1. .644
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Compton line 151"
 /db_xref="taxon:9031"
 /clone="CHEST508C24"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH103"
 /clone_lib="CSEQCHN61"
 /note="Organ: heart; Vector: pBluescript II KS(+); Site_1: EcRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended ligated to NotI adapters, digested with EcRI, size-selected, and cloned into the NotI and EcRI

RESULT 11

CB547284/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

CB547284 402 bp mRNA linear EST 01-APR-2003
 AMGNNUC:SRP2-00242-E3-A srp2 (10220) Rattus norvegicus cDNA clone
 srp2-00242-e3 5', mRNA sequence.

CB547284
 CB547284.1 GI:29431225
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 402)
 Angen EST Program.

compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 192 a 141 c 168 g 142 t 1 others
 ORIGIN

Query Match 19.4%; Score 389; DB 13; Length 644;
 Best Local Similarity 75.5%; Pred. No. 3.7e-68;
 Matches 482; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
 QY 1166 CCAGTTCTGACAGTGGAGCGTCTGCACTTCTCTGCGGAACCTTCTGGTGTGCTAC 1225
 Db 7 CGAGCAGACAGCAGTGGAGCGTCTTACATCTCTCTTCAGGAACCTTTTGGCATGGCTGC 66
 QY 1226 AGCTGAAGATGATGAATTAAGCGCGCAGCCACCTATTGGAGCGACTTTCCAGCAGTTTC 1285
 Db 67 AATTGAAGGAGGATGAATTAACAGCAGCAGCCCTATTGGTGAGATATCCCACTGTGC 126
 QY 1286 AGAAGCAGAACGATGTACATAGGGCTTCAAGAGGAATTGAAAACTAAAGAACCTGTAA 1345
 Db 127 AGAAGCAGAAATGATTTTATAGGACTTTCAAGAGGAGCTGAAAACAAAAGAACCTGTTA 186
 QY 1346 TCATGATCTCTTGGAGCTGTACGATATTTTCTGACAGCAGCCTTTGGAGGAGCTAG 1405
 Db 187 TCATGAATGACATTTGAGACTGTGCGACTCTTCTGGCAGATCAACAGTAGAGGAGCTGG 246
 QY 1406 AGAAACTCTTACCAGGAGCCAGAGAGCTGCCCTCTGAGGAGAGAGCCAGAAATGTCACTC 1465
 Db 247 AAAAGTCTATCCAGAACCAAGAGACTTATCACTGAGGAGAGGGCCAGAAATGTCACTA 306
 QY 1466 GGCTTCTAGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAAAAATTGAACCTGCAT 1525
 Db 307 AAGTTTCTCCAGAGCAGCAGATGATGTGAGAACTGAGTGGGTAAGCTAAATCTACGTT 366
 QY 1526 CCCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACTCCAGGAACCTCAAGAGG 1585
 Db 367 CTGCTGATGGGAAAAGAGATAGATGATGCTTTGAAAGACTGCAAGGTCTTCAGGAGG 426
 QY 1586 CCACGATGAGTGGACCTCAAGTGGCGCAAGTGGTGTGATCAAGGGATCTGGCAGC 1645
 Db 427 CAATGGATGAATAGACCTGAAACTCGCCAGGCTGAAGCATTCAAGGGATCTGGCAGC 486
 QY 1646 CCCTGGGCGATCTCCTCATTTGATCTCTCCAGATCACTCGAGAAAGTCAAGGCACCTTC 1705
 Db 487 CAGTGGGGGATCTGCTGATAGACTCTCTGAGATCACTTAGAAAAGTCAAGGTTTATC 546
 QY 1706 GAGGAGAAATTTGGCCCTCTGAAAGAGAGACGTGAGCCACGTCAATGACCTTGTCCGCCAGC 1765
 Db 547 GAGCAGAAATGGTGGCCCTTANAGAGAAGGTGCATCAAGTCAATGAGCTGGCTCACCGGT 606
 QY 1766 TTACCACTTGGGCATTCAGCTCTCACCGTATAACCTC 1803
 Db 607 TCGCTCCCTCGATATTCAGTTCTCCCATACACTCTC 644

```

TITLE
JOURNAL
COMMENT
Amgen Rat EST Program
Unpublished
Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00242 row: e column: 3.
Location/Qualifiers
1. 402
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="srp2-00242-e3"
/tissue type="prostate tissue"
/clone_lib="srp2 (10220)"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; rat
prostate normalized double selected poly(A+) mRNA size
fraction > 1 kb"
3BASE COUNT 83 a 111 c 88 g 119 t 1 others
ORIGIN
Query Match 16.2%; Score 324.6; DB 14; Length 402;
Best Local Similarity 90.8%; Pred. No. 3.4e-55;
Matches 345; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
2Y 1236 TGATGAATTAAGCCGCGACGACCTATTGGAGCGGACTTTCCAGCAGTTTCAGAACGAGAA 1295
2b 402 TGATGATTTAGCCGCGACGACCTATTGGCGGATTTTCCAGCAGTTTCAGAACGAGAA 343
2Y 1296 CGATGTACATAGCGCTTCAAGAGGAATTAAGAACTAAGACCTGTAATCATGAGTAC 1355
2b 342 TGATGTACATAGCGCTTCAAGAGGAATTAAGAACTAAGACCTGTAATCATGAGTAC 283
2Y 1356 TCCTGAGACTGTACGAAATATTTCTGACAGCAGCGCTTTGGAGGACTAGAGAACTCTA 1415
2b 282 TCCTGAGACTGTACGAAATATTTCTGACAGCAGCGCTTTGGAGGACTAGAGAACTCTA 223
2Y 1416 CAGAGGCCGAGAGAGCTGCTCTGAGAGAGAGCCGAGATGTCATCTGGCTTCTAGG 1475
2b 222 CCAGGAGCCGAGAGAGCTGCTCTGAGAGAGAGCTCAGATGTCTGCTGGCTCTAGG 163
2Y 1476 AAGCAGGCTGAGAGGTCATTAATCACTGAGTGGGAAATTTGAACCTGCACTCGGCTGACTG 1535
2b 162 AAGCAGGCTGAGAGGTCATTAATCACTGAGTGGGAAATTTGAACCTGCACTCGGCTGACTG 103
2Y 1536 CGAGAGAAATAGATGAGACCTTCAAGAGCTCCAGGAACTCCAGAGGCCACCGATGA 1595
2b 102 CGAGGAAATAGATGAGGCTCTTCAAGAGCTCCAGGAACTCCAGAGGCCACCGATGA 43
2Y 1596 GCTGGACCTCAAGCTGGCC 1615
2b 42 ACAGCAGCTCATCTTGGCGC 23
RESULT 12
LOCUS AK081426
DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:Cl30016K19 product:dystrophin related protein 2, full insert sequence.
ACCESSION AK081426
VERSION AK081426.1 GI:26349154
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

```

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED 11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsue, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED 11076861

4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staehli, F., Suzuki, R., Tomita, M., Wagner, J., Washio, T., Sakai, K., Okido, T., Furuno, M., Anon, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., De Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Taya, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
PUBMED 11217851

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3753)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome

COMMENT

Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

Location/Qualifiers
1..3753
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="C57BL/6J"
/db_xref="FANTOM_DB:C130016K19"
/db_xref="taxon:10090"
/clone="C130016K19"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="16 days embryo"
181..3054
/note="unnamed protein product; dystrophin related protein
2 (MGD|GI:107432, GB|U43520, evidence: BLASTN, 99%,
match=807)
putative"
/codon_start=1
/protein_id="BAC38217.1"
/db_xref="GI:26349155"
/translation="MQPLVMGQCPYTLPRCHEWHAADRFHSSSLRNTCPQPOVRAAV
TIPAPPWDGADPCLSKPLNGTVGATGPLEPSPAMNLCWNEIKKHNLRLARLEAFSD
LSGKLPQPLREIIIDWLSKOBELSAQLPLOGVALVOEKETAHAAPFEEVKSQPIYS
SVLSAQAQLSOHPFELESQSSKTSRPRIONLRFVWKOATVASLEWKLQAR
CVDQRIEHLHLHLLIQGMEELSTLQAEGVRATWEPIDGLDIFDISLPEHQAIR
LFKEEFPKQGVKLVNDLHQAISDVHLSMENSRALEQINIRWQLOQVSAERLQK
LQDAHDFGSGHSLTSVQWPERAFSPKNVPIYINHOAQTTCDMDHPKOTELYQTL
ADLNRIKFSAYRTAMKLRVOKALRLDLVLTITALEIFNBDHQAQSEHMDVVEVHC
LITALYERLEBERGLVNVPLCDMSLNMNLNVDGSRGKRALSPKTIACLCGTEV
KEKQYLPSOVANGSGCDORHLGALLHEATQVPROLGEVAAGFSGNVSPVSRCFR
STGKPLVASQFLQWNLQPSWVLAVLHVIAEOVKHOKTSCICROCPKIGPRVY
SLKFNVDICQTCITLGRASGNKLHPYIMEYITPTSSNMEDPATLLANKERSQY
FSKHPQKGLPQVLESQDSGETPASPMPHADTHSRIFASFARLAMESQNSCFN
DLSLPDSDIDBQVLLRHSSPTIDREPAFGQPCSWATESGKEKILAHLEDRNLI
LQGLRLKQWHEAAAPATVEGSAEATPDHNEELAEARLRQHSLETRMQL
EDHNKQLESQRLRELLQPPSESDNGNSGASGLASSPQSESGSPRERKQTPDTE
VADDVGSQSDVSLCLEDIMEKLHAFPSVRSSDVTANTLLAS"

polyA_signal 3731..3736
/note="putative"

polyA_site 3753
/note="putative"

BASE COUNT 927 a 1021 c 938 g 867 t

ORIGIN

Query Match 15.5%; Score 309.8; DB 11; Length 3753;
Best Local Similarity 59.1%; Pred. No. 4.2e-52;
Matches 530; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 1104 CATGAACCTCAAGTGGAGCTTGGAAAGATCTCTCAACATTAGGTCCCATTTGGA 1163
DB 414 CATGAATCTGTGTTGGAATGAATAAAGAGAGTCTCACAACTCCGCGCTCGCTAGA 473
QY 1164 ACCGAGTTCTGACAGTGGAGCGTCTGCACCTTCTCTGAGGAACCTTCTGGTGGCT 1223
DB 474 GGCCTTCTCAGACCTTCAGTGGAAACCTTCAGTCCCGCTCGAGAGATTATTGACTGGCT 533
QY 1224 ACAGCTGAAGATGATGAATTAAGCGCGCAGCACCTATTGGAGGCGACTTCCAGCAGT 1283
DB 534 CAGCAAAAGGATGAGGATGTTGAGTCACTCCCTTTCAGAGGATGTTGGCCCTGGT 593
QY 1284 TCAGAGCAGACGATGATGATAGAGGCTTCAAGAGGGAATTAAGAACTAAGAACCTGT 1343
DB 594 ACAACAGGAGAGGAGACACATGACGCTTCATGAGAGAGTCAAGTCTAAGGCCCTTA 653
QY 1344 AATCATGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGAGCTTTTGGAGAGCT 1403
DB 654 CATCTCCTCTGTGTTGAATCCGCCCGCCAGCTTTCTGTCCCGCATCATTTGAGGAATT 713

QY 1404 AGAGAACTCTACACAGGAGCCAGAGAGCTCCCTCTGAGGAGAGAGCCCAAGATGTCAAC 1463
DB 714 AGAGAGTCTCAATCTGAGAGTAAGATACCTCCCGACAGACAGGGATTCAAACCTTAG 773
QY 1464 TCGGCTTCTACAAAGCAGGCTGAGGAGTCAATCTGAGTGGGAAAAAATTAACCTGCA 1523
DB 774 TCGCTTTCTATGGAAGCAGGCAACAGTGGCAGTGAGCTGTGGGAGAGCTGACAGCCCG 833
QY 1524 CTCCTGCTGCTGGCAGAGAAAAATAGATGACACCTTGAAAGACTCCAGGAACCTCAAGA 1583
DB 834 CTGTGTAGATCAGCATCCCCACATTGACATCTCTGGAACATCTATTGGAGATCCCAAG 893
QY 1584 GGCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGATCCTTGCA 1643
DB 894 GGCATATGAGGAATTTGACAGTACTTTTGAACCAAGCAGAGGAGTCCGAGCCACGTTGGA 953
QY 1644 GCGCTGGCGATCTCTCATTTGACTCTCTCCAGATCACCCTCGAAGAGTCAAGGACT 1703
DB 954 GCGCATAGAGATCTCTTTATTGATTCCTCCAGAGCATATCCAAGCATCAAGCTATT 1013
QY 1704 TCGAGGAGAAATTTGCGCTCTGAAAGAGAAAGTGGAGTGGAGTGGAGTGGAGTGGAG 1763
DB 1014 CAAAGAGAAATTTCTCTCTGTGAAAGATGGGTTAGTGAATGATCTGGCCACCA 1073
QY 1764 GCTTACCATTTTGGGATCTAGCTCTCACTGATTAACCTGAGCATCTGGAAGACTGAA 1823
DB 1074 GCTTCCCATCTCTGATGTGCACTTGTCAATGGAGAAATTCAGGGCTCTGGAAACAGATCAA 1133
QY 1824 CACCAAGTGGAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAAGGAGCTGAGTCAAGC 1883
DB 1134 CATCCGTTGGAACAGCTCCAGGTGTCACTGGCTGAGAGGCTTAAGCACTCCAGGATGC 1193
QY 1884 CCACAGGAGCTTTGCTGAGCATCTCAGCATCTTCTTCCAGCTTGTCCAGGCTCCCTG 1943
DB 1194 CCACGGGACTTTGGGCTGGCTGAGTCTTCTCTCCACTTCTGTCCAGTTCCCTG 1253
QY 1944 GAGAGAGCATCTCGCCAAACAAAGTCCCTACTATATCAACACAGAGACTCAAAAC 2000
DB 1254 GGAAGAGCATTTTCTCCCAATAAAGTCCCTACTATCAATCAACACAGGCTCAGAC 1310
RESULT 13
BC036095 3051 bp mRNA linear HTC 04-MAR-2003
LOCUS
DEFINITION Homo sapiens, clone IMAGE:5300319, mRNA.
ACCESSION BC036095
VERSION BC036095.1 GI:23271908
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3051)
Straussberg, R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help Desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxill@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 48 Row: 0 Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4503392
This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1..3051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5300319"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/notes="Vector: pBluescript"
BASE COUNT 848 a 724 c 725 g 754 t
ORIGIN
Query Match 13.3%; Score 265.4; DB 11; Length 3051;
Best Local Similarity 57.8%; Pred. No. 3.8e-43;
Matches 473; Conservative 0; Mismatches 346; Indels 0; Gaps 0;
2Y 1105 ATGAATCTCAAGTGGAGTGAACCTTCGGAAAGAGTCTCTCAACATAGTCCCATTTGGAA 1164
DB 553 ATGAATCTGTGTGAATGAAATATAAAGAGAGTCTCAACAGCTCGCGCTCGCCTAGAG 612
2Y 1165 GCCAGTTCTGACAGTGGAGCGTCTGCACCTTTCTCGAGGAAGTCTCTGTGTGGCTA 1224
DB 613 GCCTTCTCAGACCAAGTGAAGCTTCAGTCCCTCTTCAGAGATATTAGCTGGCTC 672
2Y 1225 CAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGAGGCGACTTTCCACAGTT 1284
DB 673 AGSCAAAGAGTGAAGAGTGTGAGCTCAGCTGCCCTACAGGGGATGTGGCCCTGTG 732
2Y 1285 CAGAGCAGACATGTACATAGGCGCTTCAGAGGGATTTGAAAGTAAAGAACTGTA 1344
DB 793 ATCTATTCTGTGCTGGAGTCAGTCAAGCGCTTCTGTGCCAGCACCCTTTGAGGAGTTA 852
2Y 1405 GAGAACTTACAGAGAGCCAGAGAGTCTCTCTGAGGAGAGAGCCAGAAATCTACT 1464
DB 853 GAGGAGCCTCATTTCTGAGAGCAAGATACCTCCCGAAACAGCGGATCCAGAACTCAGC 912
2Y 1465 CGGCTTCTAGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAATTTGAACCTGCAC 1524
DB 913 CGCTTTGTATGAGAGAGGCGGCTGGCCAGTGAATCTGTGGAGAGTTGACAGCCGC 972
2Y 1525 TCCGCTGACTGGCAGAGAAAATAGATAGAGACCTTTGAAAGACTCCAGAACTTCAAGAG 1584
DB 973 TGTGTGAGCAGCAGCCGTACATTTGAGCGGACTCTGGAGCAGCTTTGGAGATTCAGGG 1032
2Y 1585 GCCACGGATGAGTGGACCTCAGCTGCCCGCAGCTGAGTGTATCAAGGATCTCTGGCAG 1644
DB 1033 GCAATGAGAGACTAAGCACTACTCTGAGCCAGCTGAGGGAGTCCGAGCCACTTGGGAG 1092
2Y 1645 CCGGTGGGGATCTCCTCTATTGACTCTCTCAAGATCACTCTCGAAAGTCAAGCACTT 1704
DB 1093 CCATTTGGGATCTCTTCTATTGATTCATCTCCAGAGCAGATCCAGGCTATTAAAGTGTT 1152
2Y 1705 CGAGGAGAAATTCGGCTCTGAAAGAGACGTGAGCCAGTCAATGACCTCTCTGCCAG 1764
DB 1153 AAAGAGAAATTTCTCCCATGAAAGATGAGTGAAGTGTGTAATGATCTGGCCCAACAA 1212
2Y 1765 CTTCACACTTTGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAC 1824
DB 1213 CTTCGCCATTTCTGATGTGACTTGTCAATGGAGATTTCCAGGCCCTTGAACAGATCAAC 1272
2Y 1825 ACCAGATGGAAGCTTCTGCGAGTGGCGCTGAGGAGCCGAGTCAAGGAGTGCATGAGCC 1884

Db 1273 GTCCGATGGAACAACTACAGGCGTCAGTTAGTGAGAGGCTTAAGCAGCTCCAGGATGCC 1332
QY 1885 CACAGGAGCTTTGGTCAGCATCTCAGCACTTTCTTCC 1923
Db 1333 CACCGGACTTTGGGCTGGTGCACAGCACTTTCTTCC 1371

RESULT 14

BC011062 2334 bp mRNA linear HTC 04-MAR-2003
LOCUS Mus musculus, Similar to utrophin, clone IMAGE:3979320, mRNA.
DEFINITION BC011062
ACCESSION BC011062
VERSION BC011062.1 GI:15029695
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2334)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gundaratne, F.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounsged, H., Kowis, C.R., Speed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 23 Row: j Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein
This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
1..2334
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3979320"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/notes="Vector: pCMV-Sport6"
BASE COUNT 698 a 512 c 638 g 486 t
ORIGIN

Query Match 13.1%; Score 261.2; DB 11; Length 2334;
Best Local Similarity 61.7%; Pred. No. 2.6e-42;
Matches 416; Conservative 0; Mismatches 258; Indels 0; Gaps 0;
QY 4 AGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTATCAACAGCTTTAGAGAGTA 63
Db 1042 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACAGATAGCGCTAGAGGAATG 1101
QY 64 TTATCGTGGCTTTCTTCTGCTGAGGACATTCGAAGCAGGAGAGATTTCTATGAT 123

Db 1102 CTGACGTGGCTGCTCCGCGGAGGACACGCTTCAGGAGCAAGATGACATTTCTGATGAT 1161
Qy 124 GTGGAAGTGGTGAAGACAGCTTTCACTACTCATGAGGGTACATGATGGATTGACAGCC 193
Db 1162 GTGGAAGTGGTGAAGACAGCTTTGCTACCATGAACTTTATGATGGAGCTGACAGCA 1221
Qy 184 CATCAGGCGCGGTGGTGAATATTCTCAATTTGGGAAGTAGCTGATGGAAACAGGAAAA 243
Db 1222 CACCAAGACAGCTGGGAGCGCTCTCGAGGTGGCAACAGCTGATGACACAGGAGCT 1281
Qy 244 TTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCTTAATTCAGATGG 303
Db 1282 CTGTCCAGAGGAGGAGAGCTTTGAGATCCAGGAACAGATGACCTTTGCTCAATGCAAGTGG 1341
Qy 304 GAATGCTCAGGCTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATGTTTAAATGAT 363
Db 1342 GAGGCGCTCCGCTGGAGAGCATGAGAGGAGCTCCCGCTGCAGGCTCTGATGGAG 1401
Qy 364 CTCAGAATCAGAACTGAAGAGTGTGAATGACTGGCTTAACAAGAAACAGAGAAAGAAC 423
Db 1402 CTGCAGAGAAACAGCTGCAGAGCTCTCAAGCTGGCTGGCGCTCCACAGAGAGCGCAT 1461
Qy 424 AGGAAATGGAAGAGAGAGCTCTTGGACCTGATCTTGAAGACCTAAAGCGCAAGTACAA 483
Db 1462 CAGAAGATGGAAGAGAGCTCCCGCTGGTGTGATGACCTGCCCTCCCTGCAGAGCTCTCAA 1521
Qy 484 CAACATAAGTGTCTTCAAGAGATCTAGAACAGAACAAAGTCAAGGTCAATTTCTCACT 543
Db 1522 GAACATAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGGTAAATTCCTTAAT 1581
Qy 544 CACATGCTGGTGTAGTGAATCTAGTGAATCTAGTGAATCTAGTGAATCTAGTGAATCT 603
Db 1582 CACATGCTGGTGTAGTGAATCTAGTGAATCTAGTGAATCTAGTGAATCTAGTGAATCT 1641
Qy 604 CAACTTAAGTATTGGAGATCGATGGGCAACATCTCTAGTGAATCTAGTGAATCTAGTGA 663
Db 1642 CAGTTACAGAACTGGTGGAGCGCTGGAGAGCTGATCCGCTGGAGTGAAGAGCTGG 1701
Qy 664 GTTCTTTACAGA 677
Db 1702 AACAGGTTGCAAGA 1715

AK034383 1541 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:9330185F14 product:dystrophin related
protein 2, full insert sequence.
AK034383
AK034383.1 GI:26329906
HTC; CAP trapper.
Mus musculus (house mouse)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakita, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, D., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Brownstein, M.J., Bult, C.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1541)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koyas, K., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 1541
/organism="Mus musculus"
/mol_type="mRNA"

```
/strain="CS7BL/6J"
/db_xref="PANTOM DB:9330185F14"
/db_xref="taxon:I0090"
/clone="9330185F14"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RKEN full-length enriched mouse cDNA library"
/dev_stages="adult"
180...1241
CDS
/note="unnamed protein product; dystrrophin related protein
2 (MGD|MG1.107432, GB|U43520, evidence: BLASTN, 99%,
match=807)
putative"
/codon_start=1
/protein_id="BAC28692.1"
/db_xref="GI:26322907"
/translation="MCPLVMGCPVTLPRCHWEHAADRFHSSSLRNTCPQPOVRAAV
TIPAPPWAGDPCLSPLKLANGTGTGATGPLESAMNLWNELKKSHNLBARLEPSD
LSGKLQPLRLRIIDWLSOKDELSAQLPQSDVALVQCEKETHAATFVEVYKSPYIS
SVLESAQPLSQHPFEELDEESSKSDTPQRIONISRFVWKOATFAATVYKSLKLTAR
CVDQRHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEH
LFKEEFSPVKGVLVNDLHQLAISDVHLSMENSRALEQINIRWKQLQVSAERLQK
LQDAHRDFGPGSQHFLSSM"
polya_signal
1521...1526
/note="putative"
polya_site
1541
/note="putative"
BASE COUNT 383 a 406 c 396 g 356 t
RIGIN

Query Match 12.9%; Score 257.8; DB 11; Length 1541;
Best Local Similarity 57.1%; Pred. No. 3.2e-41;
Matches 469; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

Y 1104 CATGAACATTCAGTGGAGTGAACATTCGGAAAGAGTCTCTCAACATTAGTCCCATTTGGA 1163
b 413 CATGAATCTGTGTTGGAATGAATAAAGAGAGTCTCAACACCTCGCGCTCGCTAGA 472
Y 1164 AGCCAGTTCTGACAGTGGAGCGTCTGCACCTTCTCTGCAGGAACCTCTGTGTGCT 1223
b 473 GGCCTTCTCAGACCTCAGTGGAAACATTCAGCTCCGCTCCGAGAGATTATTGACTGCT 532
Y 1224 ACAGCTGAAGATGATGAATTAAGCCGCGAGCAGCTATTGGAGCGCAGTTTCCAGCAGT 1283
b 533 CAGCCAAAGGATGAGGATGTCAGCTCAGCTCCCTTGCAGGGGATGTGGCCCTGCT 592
Y 1284 TCAGAAGCAGAACGATGTACATGAGGCTTCAAGAGGAAATGAAAACCTGTT 1343
b 593 ACAACAGGAGAGGAGACACATGAGCCTTCATGAGAGAGTCAAGTCTAAGGGCCCTA 652
Y 1344 ATCATGAGTACTCTTGAGCTGTAGCAATATTCTGACAGCAGCAGCTTTGGAAGGACT 1403
b 653 CATCTCTCTGTGCTTGAATCCGCCCGAGCTTTCTGCCAGCATCTTGAAGATT 712
Y 1404 AGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTAC 1463
b 713 AGAGAGTCTCAATCTGAGAGTAAAGATACCTCCCCAGACAGCGGATTCAAAACCTAG 772
Y 1464 TCGGCTTCTACGAAGCAGGCTGAGAGGTCATCTACTAGTGGGAAAATTTGAACCTGCA 1523
b 773 TCGCTTTGTATGGAAGCAGGCAACAGTGGCCAGTGAGCTGTGGGAGAGAGCTGCAGGCCG 832
Y 1524 CTCCTGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAGAGACTCCAGGAACCTCAAGA 1583
b 833 CTGTGTAGATCAGCATCGCCCATTTGAGCATCTCTGGAACATCTATTGGAGATCCAAGG 892
Y 1584 GGCACGGATGAGCTGGACCTCAAGCTCGGCGAGCTGAGTGTATCAAGGATCTCTGGCA 1643
b 893 GGCAATGGAGGAATTTGAGCAGTACTTTGACCCAGCAGAGGAGTCCAGGCCAGCTGGGA 952
Y 1644 GCCCGTGGGCGATCTCCCTATTGATCTCTCTCCAGATCACCTCGAGAAAGTCAAGGACT 1703
b 953 GCCCATAGGAGATCTCTTTATTGATTCCTTCCAGGCAATATCCAGCCATCAAGCTATT 1012
```

Search completed: February 2, 2004, 06:03:07
Job time : 3755.42 secs

```
QY 1704 TCAGAGAGAAATTGCGCCTCTGAAGAGAAACGTGAGCCACGTCAATGACCTTGCTCGCCA 1763
Db 1013 CAAGAGAGATTCTCTCCTGTGAAGATGGGGTGAAGTTAGTGAATGATCTGCCCCACCA 1072
QY 1764 GCTTACACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTGAA 1823
Db 1073 GCTTGGCATCTCTGATGTGCACTTGTCAATGGAGAATTCAGGGCTCTGGAACAGATCAA 1132
QY 1824 CACCAGATGGAAGCTTCTCAGGTGGCGCTCGAGACCGAGTCAGGAGCTGCATCAAGC 1883
Db 1133 CATCCGTTGGAACAGCTCCAGTGTCACTGTGCTGAGAGGCTTAGCAACTCCAGGATGC 1192
QY 1884 CCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCA 1924
Db 1193 CCACCGGAGCTTTGGGCTTGGGCTTGGGCTCAGCACTTCTCTCCA 1233
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

M nucleic - nucleic search, using sw model

run on: February 1, 2004, 23:51:42 ; Search time 655.123 Seconds
(without alignments)
11131.886 Million cell updates/sec

File: US-09-845-416-6_COPY_1000_3000

Perfect score: 2001

Sequence: 1 ggcattcatgatggagag.....tcaaccagagactcaaca 2001

Scoring table: IDENTITY NUC

Gapop 10_0, Gapext 1.0

Searched: 2434939 seqs, 182278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA.*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2001	100.0	3999	13	US-09-845-416-6
2	2001	100.0	4966	13	US-09-845-416-28
3	2001	100.0	4990	13	US-09-845-416-34
4	1709	85.4	3858	13	US-09-845-416-9
5	1709	85.4	4825	13	US-09-845-416-29
6	1709	85.4	4848	13	US-09-845-416-35
7	1709	85.4	5060	13	US-09-845-416-36
8	1662.8	83.1	4182	13	US-09-845-416-2
9	1662.8	83.1	5149	13	US-09-845-416-27
10	1283	64.1	5462	13	US-10-149-736-41
11	1182.6	59.1	8689	13	US-10-149-736-42
12	1182.6	59.1	11058	13	US-09-845-416-1
13	1182.6	59.1	11443	13	US-10-149-736-44
14	1182.6	59.1	12057	13	US-10-149-736-47
15	1182.6	59.1	13957	10	US-09-782-378A-22

```
16 1182.6 59.1 13957 10 US-09-880-107-2284
17 1182.6 59.1 13957 13 US-10-149-736-1
18 1182.6 59.1 14082 13 US-10-341-434-108
19 1182 59.1 2189 13 US-09-845-416-4
20 1182 59.1 3531 13 US-09-845-416-10
21 1182 59.1 4498 13 US-09-845-416-30
22 1180 59.0 5339 13 US-10-149-736-40
23 1013 50.6 3510 13 US-09-845-416-12
24 1013 50.6 4476 13 US-09-845-416-31
25 1004 50.2 13815 13 US-10-149-736-2
26 835 41.7 1821 13 US-09-845-416-13
27 727 36.3 5417 13 US-10-149-736-39
28 678.4 33.9 4414 13 US-09-845-416-32
29 678 33.9 1991 13 US-09-845-416-3
30 677 33.8 1667 13 US-09-845-416-7
31 666.4 33.3 3446 13 US-09-845-416-14
32 448 22.4 1434 13 US-09-845-416-15
33 408.2 20.4 11096 13 US-10-149-736-4
34 397 19.8 10302 10 US-09-782-378A-23
35 387 19.3 10302 13 US-10-149-736-3
36 387 19.3 387 13 US-10-149-736-32
37 385 19.2 16531 13 US-10-101-510-667
38 350 17.5 1340 13 US-09-845-416-11
39 348 17.4 348 13 US-10-149-736-31
40 330 16.5 333 13 US-10-149-736-9
41 327 16.3 327 13 US-10-149-736-8
42 324 16.2 324 13 US-10-149-736-33
43 160.2 8.0 256 9 US-09-864-761-21956
44 153.4 7.7 467 9 US-09-864-761-11083
45 151 7.5 151 9 US-09-864-761-27715
```

ALIGNMENTS

RESULT 1

US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF

; FILE REFERENCE: DB1142

; CURRENT APPLICATION NUMBER: US/09/845, 416

; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-845-416-6

```
Query Match      100.0%; Score 2001; DB 13; Length 3999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  GCGAGTTCATGTATGAGAGAGTGAAGTAAACCTGACCGTTATCAACACAGCTTAGAAGAA 60
Db      1000 GCGAGTTCATGTATGAGAGAGTGAAGTAAACCTGACCGTTATCAACACAGCTTAGAAGAA 1059

Qy      61  GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCCAAGCAGGAGAGATTTCCTAAT 120
Db      1060 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCCAAGCAGGAGAGATTTCCTAAT 1119

Qy      121 GATGTGAAGTGTGAAGACCAAGTTTCATCTATCATGAGGGGTACATGATGATTTGACA 180
Db      1120 GATGTGAAGTGTGAAGACCAAGTTTCATCTATCATGAGGGGTACATGATGATTTGACA 1179

Qy      181 GCCCATCAGGCGCGGTGGTGTATTCTACATTGGGAAGTAAGCTGATTGGAACAGGA 240
```

Db 1180 GCGCATCAGGCGCGGTTGGTATATTTCTACAAATGGGAAGTACCTGATTTGGAAACAGGA 1239
QY 241 AAATTATCAGAAAGATGAGAAATCTGAAGTACAGAGGAGATGAATCTCTTAATTCAGA 300
Db 1240 AAATTATCAGAAAGATGAGAAATCTGAAGTACAGAGGAGATGAATCTCTTAATTCAGA 1299
QY 301 TCGGAATGCTTCAGGATAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 360
Db 1300 TCGGAATGCTTCAGGATAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 1359
QY 361 GATCTCCAGATCAGAAATCTGAAGATGATGAATGATGAATGATGAATGATGAATGATGAAT 420
Db 1360 GATCTCCAGATCAGAAATCTGAAGATGATGAATGATGAATGATGAATGATGAATGATGAAT 1419
QY 421 ACAAGGAAATCGAGGAGAGAGCTCTTGACCTGATCTTGAGACCTTAAAGCGCAAGTA 480
Db 1420 ACAAGGAAATCGAGGAGAGAGCTCTTGACCTGATCTTGAGACCTTAAAGCGCAAGTA 1479
QY 481 CAACCAATAGGTCCTTCAAGAGATCTAGAACCAAGATCAAGGTCATCTCTC 540
Db 1480 CAACCAATAGGTCCTTCAAGAGATCTAGAACCAAGATCAAGGTCATCTCTC 1539
QY 541 ACTCATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
Db 1540 ACTCATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1599
QY 601 GAAACCAATAGGTCCTTCAAGAGATCGATGGGCAACATCTGTAGATGACAGAGAGCGC 660
Db 1600 GAAACCAATAGGTCCTTCAAGAGATCGATGGGCAACATCTGTAGATGACAGAGAGCGC 1659
QY 661 TGGTTCCTTTTACAGACAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 720
Db 1660 TGGTTCCTTTTACAGACAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 1719
QY 721 CTTACTCAGACTGTACTCTGTGTGACACAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 1720 CTTACTCAGACTGTACTCTGTGTGACACAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1779
QY 781 AAACATAGAAATGCCATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 1780 AAACATAGAAATGCCATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1839
QY 841 TTCCCTCCTGACCTGAAAGTTCTTCTGCTGCTTACAGAGCTGAAACAACTGCCAAT 900
Db 1840 TTCCCTCCTGACCTGAAAGTTCTTCTGCTGCTTACAGAGCTGAAACAACTGCCAAT 1899
QY 901 GTCCTACAGGATGCTACCCGTAAAGAAAGCTCTCTAGAGACTCCAAAGGAGTAAAGAG 960
Db 1900 GTCCTACAGGATGCTACCCGTAAAGAAAGCTCTCTAGAGACTCCAAAGGAGTAAAGAG 1959
QY 961 CTGATGAAACATGGCAGACCTCCAGGTGAAATGAACTCAGCTCAGACAGATGTTTATCAC 1020
Db 1960 CTGATGAAACATGGCAGACCTCCAGGTGAAATGAACTCAGCTCAGACAGATGTTTATCAC 2019
QY 1021 AACCTGGATGAAACAGCCCAAAATCTCTGAGATCCCTGGAGGTTCCGATGATGATGATGATG 1080
Db 2020 AACCTGGATGAAACAGCCCAAAATCTCTGAGATCCCTGGAGGTTCCGATGATGATGATGATG 2079
QY 1081 CTGTTTCAAGAGCTTTGGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1140
Db 2080 CTGTTTCAAGAGCTTTGGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 2139
QY 1141 CTCACATTTAGGTCCTTCAAGGAGCTCTCTGACCTGAGGAGCTCTGACCTTCT 1200
Db 2140 CTCACATTTAGGTCCTTCAAGGAGCTCTCTGACCTGAGGAGCTCTGACCTTCT 2199
QY 1201 CTGACGAACTTCTGTGTGGTCTACAGCTGAAAGATGATGAATTAAGCGGCGAGGACCT 1260
Db 2200 CTGACGAACTTCTGTGTGGTCTACAGCTGAAAGATGATGAATTAAGCGGCGAGGACCT 2259
QY 1261 ATTGGAGGCGACTTCCAGCTTTCAGAGCAGACGATGTACATAGGCTTCAAGAGG 1320

Db 2260 ATTGGAGGCGACTTTCCAGCAGTTTCAAGCAGAACGATGTACATAGGCGCTTCAAGAGG 2319
QY 1321 GAATTGAAACCTAAAGAACCTCTGATCATGATGATCTCTTGGAGACTGTACGAATATTTCTG 1380
Db 2320 GAATTGAAACCTAAAGAACCTCTGATCATGATGATCTCTTGGAGACTGTACGAATATTTCTG 2379
QY 1381 ACAGAGCAGCCTTTTGAAGGACTAGAGAAATCTCTACAGGAGCCCGAGAGAGCTGCTCTCT 1440
Db 2380 ACAGAGCAGCCTTTTGAAGGACTAGAGAAATCTCTACAGGAGCCCGAGAGAGCTGCTCTCT 2439
QY 1441 GAGGAGAGAGCCCGAGAGTGTCTACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACT 1500
Db 2440 GAGGAGAGAGCCCGAGAGTGTCTACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACT 2499
QY 1501 GAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTCGGAGAGAAAAATAGATGAGACCTTT 1560
Db 2500 GAGTGGGAAAAATTTGAACCTGCACCTCCGCTGACTCGGAGAGAAAAATAGATGAGACCTTT 2559
QY 1561 GAAAGACTCCAGAACTTCAAGAGGCCCGGATGAGCTGGACCTCAAGCTGCGCAAGCT 1620
Db 2560 GAAAGACTCCAGAACTTCAAGAGGCCCGGATGAGCTGGACCTCAAGCTGCGCAAGCT 2619
QY 1621 GAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGAT 1680
Db 2620 GAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGAT 2679
QY 1681 CACCTCGAGAAAGTCAAGCACTTCGAGGAGAAATGGCCCTCTGAAAGAGACGTGAGC 1740
Db 2680 CACCTCGAGAAAGTCAAGCACTTCGAGGAGAAATGGCCCTCTGAAAGAGACGTGAGC 2739
QY 1741 CAGGTCAATGACTGTCTGCGCAGCTTACACTTTTGGGCAATTCAGCTCTCACCGTATAAC 1800
Db 2740 CAGGTCAATGACTGTCTGCGCAGCTTACACTTTTGGGCAATTCAGCTCTCACCGTATAAC 2799
QY 1801 CTCAGCACTCTGAAAGACCTGAACACACAGATGGAAGTCTGAGGTGGCGCTGAGAGAC 1860
Db 2800 CTCAGCACTCTGAAAGACCTGAACACACAGATGGAAGTCTGAGGTGGCGCTGAGAGAC 2859
QY 1861 CGAGTCAGCAGCTGATGAGCCAGCCAGCAGGAGCTTTGGTCCAGCATCTCAGCACTTTCT 1920
Db 2860 CGAGTCAGCAGCTGATGAGCCAGCCAGCAGGAGCTTTGGTCCAGCATCTCAGCACTTTCT 2919
QY 1921 TCCACCTCTGTCCAGGCTCTCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGGCCCTACTAT 1980
Db 2920 TCCACCTCTGTCCAGGCTCTCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGGCCCTACTAT 2979
QY 1981 ATCAACCCAGAGACTCAACA 2001
Db 2980 ATCAACCCAGAGACTCAACA 3000

RESULT 2
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28
Query Match 100.0%; Score 2001; DB 13; Length 4966;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGAGTTCATGATGAGAGTGAAGTAAACCTCGGCGGTATCAACACGTTTAGAAGAA 60
Db 1757 GCGAGTTCATGATGAGAGTGAAGTAAACCTCGGCGGTATCAACACGTTTAGAAGAA 1816

QY 61 GTATTATCGTGGCTCTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 120
Db 1817 GTATTATCGTGGCTCTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT 1876

QY 121 GATGTGAAGTGTGAAAGACCAAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 180
Db 1877 GATGTGAAGTGTGAAAGACCAAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 1936

QY 181 GCCCATCAGGCGCGGGTGTGTAATATTTCAATTCGGAAGTAAGCTGATTCGAACAGGA 240
Db 1937 GCCCATCAGGCGCGGGTGTGTAATATTTCAATTCGGAAGTAAGCTGATTCGAACAGGA 1996

QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGAGATGATCTCCTAAATTCAGA 300
Db 1997 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGAGATGATCTCCTAAATTCAGA 2056

QY 301 TGGGAATGCGCTCAGGTPAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAATG 360
Db 2057 TGGGAATGCGCTCAGGTPAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAATG 2116

QY 361 GATCTCCAGAAATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGA 420
Db 2117 GATCTCCAGAAATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAGA 2176

QY 421 ACAAGGAAATGAGGAGAGAGCTCTTGACCTGATCTGGAAGACCTAAACGCCAAGTA 480
Db 2177 ACAAGGAAATGAGGAGAGAGCTCTTGACCTGATCTGGAAGACCTAAACGCCAAGTA 2236

QY 481 CAACAAATTAAGGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAACCGC 540
Db 2237 CAACAAATTAAGGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAACCGC 2296

QY 541 ACTCAGATGCTGAGTGTGATGATGATCTAGTGGAGATCACGCAACTGCTGTTTGGAA 600
Db 2297 ACTCAGATGCTGAGTGTGATGATGATCTAGTGGAGATCACGCAACTGCTGTTTGGAA 2356

QY 601 GAACAACTTAAGGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAACCGC 660
Db 2357 GAACAACTTAAGGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAAACCGC 2416

QY 661 TGGGTTCTTTTACAAGACCAAGCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCT 720
Db 2417 TGGGTTCTTTTACAAGACCAAGCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCT 2476

QY 721 CCTACTCAGACTGTTACTCTGTTGACACACCTGTTGTTACTTAAGGAAACTGCCATCTCC 780
Db 2477 CCTACTCAGACTGTTACTCTGTTGACACACCTGTTGTTACTTAAGGAAACTGCCATCTCC 2536

QY 781 AAACCTAGAAATGGCATCTTCTTCTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
Db 2537 AAACCTAGAAATGGCATCTTCTTCTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 2596

QY 841 TTCCCTCCGACCTGGAAGATTTCTTCCCTGCTGCTTACAGAGCTGAACCACTGCGCAAT 900
Db 2597 TTCCCTCCGACCTGGAAGATTTCTTCCCTGCTGCTTACAGAGCTGAACCACTGCGCAAT 2656

QY 901 GTCTCAGAGATGCTACCGTGAAGAAAGGCTCCTAGAGACTCCAGAGGAGTAAAGAG 960
Db 2657 GTCTCAGAGATGCTACCGTGAAGAAAGGCTCCTAGAGACTCCAGAGGAGTAAAGAG 2716

QY 961 CTGATGAAACAAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTATCAC 1020
Db 2717 CTGATGAAACAAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTATCAC 2776

QY 1021 AACCTGATGAACACAGCCAAAATCTCTGAGATCCCTGGAAGTTCCGATGATGAGTCT 1080

Db 2777 AACCTGATGAACACAGCCAAAATCTCTGAGATCCCTGGAAGTTCCTGATGATGACGTC 2836

QY 1081 CTGTTACAAAGACGCTTTTGGATAACATGAATCTCAAGTGGAGTGAACCTTCGAAAAAGTCT 1140

Db 2837 CTGTTACAAAGACGCTTTTGGATAACATGAATCTCAAGTGGAGTGAACCTTCGAAAAAGTCT 2896

QY 1141 CTCAACTTAAGTCCCAATTTGGAAGCCAGTTCCTGAACAGTGGAGCGTCTGCACTTTCT 1200

Db 2897 CTCAACTTAAGTCCCAATTTGGAAGCCAGTTCCTGAACAGTGGAGCGTCTGCACTTTCT 2956

QY 1201 CTGAGAGAACTTCTGCTGCTGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGCACT 1260

Db 2957 CTGAGAGAACTTCTGCTGCTGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGCACT 3016

QY 1261 ATTGAGAGCGCACTTTCCAGCAGTTTCAAGCAGAACTGATACATAGGGCTTCAAGAGG 1320

Db 3017 ATTGAGAGCGCACTTTCCAGCAGTTTCAAGCAGAACTGATACATAGGGCTTCAAGAGG 3076

QY 1321 GAATGAAACTAAGAACCTGTAATCATGAGTACTCTTGGAGACTGTACGAATATTTCTG 1380

Db 3077 GAATGAAACTAAGAACCTGTAATCATGAGTACTCTTGGAGACTGTACGAATATTTCTG 3136

QY 1381 ACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTTACAGAGGCCAGAGAGCTGCCTCT 1440

Db 3137 ACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTTACAGAGGCCAGAGAGCTGCCTCT 3196

QY 1441 GAGGAGAGCGCCAGAAATGTCTCGGCTTCTAGAAAGCAGGCTGAGAGGTTCAATCT 1500

Db 3197 GAGGAGAGCGCCAGAAATGTCTCGGCTTCTAGAAAGCAGGCTGAGAGGTTCAATCT 3256

QY 1501 GAGTGGGAAATTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560

Db 3257 GAGTGGGAAATTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3316

QY 1561 GAAAGACTTCCAGGAACTTCAAGAGGCCACCGATGAGTGAAGCTTCAAGCTGCGCCAGCT 1620

Db 3317 GAAAGACTTCCAGGAACTTCAAGAGGCCACCGATGAGTGAAGCTTCAAGCTGCGCCAGCT 3376

QY 1621 GAGGTGATCAAGGATCTCTGCGAGCCGCTGCGAGCTCTCTCTCTCTCTCTCTCTCTCT 1680

Db 3377 GAGGTGATCAAGGATCTCTGCGAGCCGCTGCGAGCTCTCTCTCTCTCTCTCTCTCTCT 3436

QY 1681 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACTGAGC 1740

Db 3437 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACTGAGC 3496

QY 1741 CAGCTCAATGACCTTGTCTGCGAGCTTACCACTTTGGGCAATTCAGCTCTCAGCTATAAC 1800

Db 3497 CAGCTCAATGACCTTGTCTGCGAGCTTACCACTTTGGGCAATTCAGCTCTCAGCTATAAC 3556

QY 1801 CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTTGCGCTCGAGGAC 1860

Db 3557 CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTTGCGCTCGAGGAC 3616

QY 1861 CGAGTCAGGAGCTGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTT 1920

Db 3617 CGAGTCAGGAGCTGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTT 3676

QY 1921 TCCAGCTCTGCTCAGGCTCTGCTGGAAGAGGCACTCTCGCCAAACAAAGTCCCTACTAT 1980

Db 3677 TCCAGCTCTGCTCAGGCTCTGCTGGAAGAGGCACTCTCGCCAAACAAAGTCCCTACTAT 3736

QY 1981 ATCAACCAAGAGCTCAAAACA 2001

Db 3737 ATCAACCAAGAGCTCAAAACA 3757

RESULT 3

US-09-845-416-34

; Sequence 34, Application US/09845416

; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO

1681 CACCTCGAGAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGABAGAGACGTCGAGC 1740
2539 CACCTCGAGAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGABAGAGACGTCGAGC 2598
1741 CACGTCATGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 1800
2599 CACGTCATGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAC 2658
1801 CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGAC 1860
2659 CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGAC 2718
1861 CGAGTCAGGCACTGCAATGAAGCCCAACAGGCACTTTGGTCAGCATCTCAGCACTTTCTT 1920
2719 CGAGTCAGGCACTGCAATGAAGCCCAACAGGCACTTTGGTCAGCATCTCAGCACTTTCTT 2778
1921 TCCAGCTCTGTCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 1980
2779 TCCAGCTCTGTCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 2838
1981 ATCAACACGAGACTCAAAACA 2001
2839 ATCAACACGAGACTCAAAACA 2859

RESULT 5
US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845, 416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200, 777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-29

Query Match 85.4%; Score 1709; DB 13; Length 4825;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;
1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAAGAA 60
1757 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAAGAA 1816
61 GTATTATCTGCTCTTCTCTGAGACACATTCGAAGACACAGGAGAGATTTCTAAT 120
1817 GTATTATCTGCTCTTCTCTGAGACACATTCGAAGACACAGGAGAGATTTCTAAT 1876
121 GATGTGGAAGTGGTGAAGACCGAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 180
1877 GATGTGGAAGTGGTGAAGACCGAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 1936
181 GCCCATCAGGCGCGGTGGTGAATATCTTACAAATGGGAAGTGAAGCTCATTTGGAACAGGA 240
1937 GCCCATCAGGCGCGGTGGTGAATATCTTACAAATGGGAAGTGAAGCTCATTTGGAACAGGA 1996
241 AAATTATCAGAAGTGAAGAACTGAGTACAGGACGAGATGAATCTCTAAATTCAGA 300
1997 AAATTATCAGAAGTGAAGAACTGAGTACAGGACGAGATGAATCTCTAAATTCAGA 2056
301 TGGGAATGCTCAGGCTAGCTAGCTAGTGAAGAACTTACAGGAGCCCAAGAGAGCTGCTCT 360

2057 TGGGAATGCTCAGGCTAGCTAGCATGCGAAACAAAGCAATTTACATAGAGTTTAAATG 2116
361 GATCTCCAGAACTCAAACTGAAGAGTTGAATGATCGCTAAACAAACACAGAGAAAGA 420
2117 GATCTCCAGAACTCAAACTGAAGAGTTGAATGATCGCTAAACAAACACAGAGAAAGA 2176
421 ACAAGGAAATGGAGAAAGAGCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCGCAAGTA 480
2177 ACAAGGAAATGGAGAAAGAGCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCGCAAGTA 2236
481 CAAACACATTAAGTCTCTCAAGAGATCTAGAACAGAACTAGTCAAGGTCATTTCTCTC 540
2237 CAAACACATTAAGTCTCTCAAGAGATCTAGAACAGAACTAGTCAAGGTCATTTCTCTC 2296
541 ACTCATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
2297 ACTCATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2356
601 GAAACAACTTAAGTATTTGGAGATCGATGGGCAACATCTGTAGTGGACAGAGACCGC 660
2357 GAAACAACTTAAGTATTTGGAGATCGATGGGCAACATCTGTAGTGGACAGAGACCGC 2416
661 TGGGTTCTTTTACAAAGACCGCTCAGCTAGCTCTCTGGACTGACCATTTATTGGAGCTCT 720
2417 TGGGTTCTTTTACAAAGACCGCTCAGCTAGCTCTCTGGACTGACCATTTATTGGAGCTCT 2434
721 CCTACTCAGACTGTTACTCTCTGGTGACAAACCTGTGGTTACTTAAGGAAACTGCCCATCTCC 780
2435 ----- 2434
781 AAACAGAAATGCCATCTTCTTGTATGTTGGAGGTACCTACTCATAGATTACTGCAACAG 840
2435 ----- ACTCATAGATTACTGCAACAG 2455
841 TTCCCTCTGACCTGGAAGAGTTTCTTCTGGCTTACAGAGCTTGAACAACTGCCAAT 900
2456 TTCCCTCTGACCTGGAAGAGTTTCTTCTGGCTTACAGAGCTTGAACAACTGCCAAT 2515
901 GTCTTACAGATGCTTACCGTGAAGAAAGGCTCTTGAAGAGCTTCAAGGAGTAAAGAG 960
2516 GTCTTACAGATGCTTACCGTGAAGAAAGGCTCTTGAAGAGCTTCAAGGAGTAAAGAG 2575
961 CTGATGAAACAAATGGCAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC 1020
2576 CTGATGAAACAAATGGCAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAC 2635
1021 AACCTGGATGAAACAGCCAAACAAATCTGAGATCCCTGGAGGTTCCGATGATCGATC 1080
2636 AACCTGGATGAAACAGCCAAACAAATCTGAGATCCCTGGAGGTTCCGATGATCGATC 2695
1081 CTGTTACAAAGACGTTTGGATTAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAGTCT 1140
2696 CTGTTACAAAGACGTTTGGATTAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAGAGTCT 2755
1141 CTCACATTTAGTTCCTCAATTTGAAGCCAGTCTTGAAGCCAGTCTGACCTTTCT 1200
2756 CTCACATTTAGTTCCTCAATTTGAAGCCAGTCTTGAAGCCAGTCTGACCTTTCT 2815
1201 CTCAGGAACTTCTGCTGGCTCAGAGCTGAAGATGATGAATTAAGCCGCGCAGCACCT 1260
2816 CTCAGGAACTTCTGCTGGCTCAGAGCTGAAGATGATGAATTAAGCCGCGCAGCACCT 2875
1261 ATTGGAGGAGCTTTCCAGCAGTTCAGAGCAGAAACGATGTACATAGGCGCTTCAAGAGG 1320
2876 ATTGGAGGAGCTTTCCAGCAGTTCAGAGCAGAAACGATGTACATAGGCGCTTCAAGAGG 2935
1321 GAATTGAACAACTTAAGAACTCTGATCATGAGTCTTGTAGACTGTGAGATATTCTG 1380
2936 GAATTGAACAACTTAAGAACTCTGATCATGAGTCTTGTAGACTGTGAGATATTCTG 2995
1381 ACAGAGCAGCTTTGGAAGAGCTAGAGAACTCTACAGAGCCCAAGAGAGCTGCTCT 1440
2996 ACAGAGCAGCTTTGGAAGAGCTAGAGAACTCTACAGAGCCCAAGAGAGCTGCTCT 3055

QY 1441 GAGGAGAGAGCCCAAGATGTCACTCGGCTTTACGAAAGCAGGCTGAGGAGTCAATACT 1500
Db 3056 GAGGAGAGAGCCCAAGATGTCACTCGGCTTTACGAAAGCAGGCTGAGGAGTCAATACT 3115
QY 1501 GAGTGGGAAATTAAGCTGCACTCGCTGACTGCGAGAGAAATAGATGAGCCCTT 1560
Db 3116 GAGTGGGAAATTAAGCTGCACTCGCTGACTGCGAGAGAAATAGATGAGCCCTT 3175
QY 1561 GAAAGACTCCAGGAACCTCAAGAGGCCACCGATGAGCTGAGCTCAAGCTGCGCCCAAGCT 1620
Db 3176 GAAAGACTCCAGGAACCTCAAGAGGCCACCGATGAGCTGAGCTCAAGCTGCGCCCAAGCT 3235
QY 1621 GAGTGATCAAGGATGCTGCGAGCCGCTGGGCGATCTCTCATTTGATCTCTTCCAGAT 1680
Db 3236 GAGTGATCAAGGATGCTGCGAGCCGCTGGGCGATCTCTCATTTGATCTCTTCCAGAT 3295
QY 1681 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGCTGAGC 1740
Db 3296 CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGCTGAGC 3355
QY 1741 CACGTCATGACCTTGTGCGGCACTTACCACTTTGGGCAATTCAGTCTTCCCGTATAAC 1800
Db 3356 CACGTCATGACCTTGTGCGGCACTTACCACTTTGGGCAATTCAGTCTTCCCGTATAAC 3415
QY 1801 CTCAGCACTCTGGAGACCTGACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGAC 1860
Db 3416 CTCAGCACTCTGGAGACCTGACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGAC 3475
QY 1861 CGAGTCAGGAGCTGATGAGCCACAGGAGCTTTGTCGAGCATCTCAGCACTTTCTT 1920
Db 3476 CGAGTCAGGAGCTGATGAGCCACAGGAGCTTTGTCGAGCATCTCAGCACTTTCTT 3535
QY 1921 TCACGCTGTCCAGGTGCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 1980
Db 3536 TCACGCTGTCCAGGTGCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 3595
QY 1981 ATCAACACAGAGCTCAACA 2001
Db 3596 ATCAACACAGAGCTCAACA 3616

RESULT 6
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 85.4%; Score 1709; DB 13; Length 4848;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GCGAGTTCATTGATGGAGATGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1780 GCGAGTTCATTGATGGAGATGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1839
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACATTTGAAAGAGAGAGATTTCTTAAT 120

Db 1840 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAAAGCAGGAGATTTCTTAAT 1899
QY 121 GATGTGGAAGTGGTGAAGACACAGTTTCATCTACTCATGAGGGGTACATGATGATTTGACA 180
Db 1900 GATGTGGAAGTGGTGAAGACACAGTTTCATCTACTCATGAGGGGTACATGATGATTTGACA 1959
QY 181 GCCCATCAGGCGCGGTTGGTAAATTTCTACAATTTGGGAAGTAAAGCTGATTTGGACAGGA 240
Db 1960 GCCCATCAGGCGCGGTTGGTAAATTTCTACAATTTGGGAAGTAAAGCTGATTTGGACAGGA 2019
QY 241 AAATTTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA 300
Db 2020 AAATTTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA 2079
QY 301 TGGGAATGCCCTCAGGCTAGCTAGCATGAGGAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2080 TGGGAATGCCCTCAGGCTAGCTAGCATGAGGAAACAAAGCAATTTACATAGAGTTTAAATG 2139
QY 361 GATCTCCAGAAATCAGAAACTGAAGAGTTGAATGACTGGCTAAACAAACACAGAGAAAGA 420
Db 2140 GATCTCCAGAAATCAGAAACTGAAGAGTTGAATGACTGGCTAAACAAACACAGAGAAAGA 2199
QY 421 ACAAGGAAATTTGGAGGAGAGCCCTTTGGACCTGATCTTTGAAGACCTTAAACGGCAAGTA 480
Db 2200 ACAAGGAAATTTGGAGGAGAGCCCTTTGGACCTGATCTTTGAAGACCTTAAACGGCAAGTA 2259
QY 481 CAACACATTAAGTGTCTTCAAGAGAGTCTAGAACAGAACAGTCAAGTCAATTTCTCTC 540
Db 2260 CAACACATTAAGTGTCTTCAAGAGAGTCTAGAACAGAACAGTCAAGTCAATTTCTCTC 2319
QY 541 ACTCACATGT 600
Db 2320 ACTCACATGT 2379
QY 601 GAACAACTTAAGT 660
Db 2380 GAACAACTTAAGT 2439
QY 661 TGGGTTCTTTTACAAAGACCGCTGACCTAGCTCTCTGAGCTGACCACTATTGTGAGCTCT 720
Db 2440 TGGGTTCTTTTACAAAGAC----- 2457
QY 721 CTTACTCAGACTGTCTCTGTGTGACAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 2458 ----- 2457
QY 781 AAACTAGAAATGCCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 2458 -----ACTCATGTATTTACTGTCAACAG 2478
QY 841 TTCCCTCTGGACCTGGAAAGTCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 2479 TTCCCTCTGGACCTGGAAAGTCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2538
QY 901 GTCTTACAGGATGTACCCGTGAAGAAAGGCTCTAGNAGACTCCAGGGAGTAAAGAG 960
Db 2539 GTCTTACAGGATGTACCCGTGAAGAAAGGCTCTAGNAGACTCCAGGGAGTAAAGAG 2598
QY 961 CTGATGAAACAAATGGCAAGACTCTCAAGGTGAATTTGAAGCTCAACAGATGTTTATCAC 1020
Db 2599 CTGATGAAACAAATGGCAAGACTCTCAAGGTGAATTTGAAGCTCAACAGATGTTTATCAC 2658
QY 1021 AACCTGGATGAAACAGCCAAATAATCTGTGATGCTCTGTGAAAGTTCGGATGTGAGTGC 1080
Db 2659 AACCTGGATGAAACAGCCAAATAATCTGTGATGCTCTGTGAAAGTTCGGATGTGAGTGC 2718
QY 1081 CTGTTTACAAAGAGCTTTGGATAACATGAATTTCAAGTGGAGTGAATTTGCGGAAAGAGTCT 1140
Db 2719 CTGTTTACAAAGAGCTTTGGATAACATGAATTTCAAGTGGAGTGAATTTGCGGAAAGAGTCT 2778
QY 1141 CTCACATTTAGTCTCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 1200
Db 2779 CTCACATTTAGTCTCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTTTCT 2838


```
QY 1201 CTGAGGAACTTCTGGTGTGGCTACAGTGAAGATGATGAATTAAGCCGCGAGGCACCT 1260
DB 2839 CTGAGGAACTTCTGGTGTGGCTACAGTGAAGATGATGAATTAAGCCGCGAGGCACCT 2898
QY 1261 ATTGGAGGCGACTTCCAGAGAGTTCAGAGCAGAGCAGATGATACATAGGGGCTTCAAGAGG 1320
DB 2899 ATTGGAGGCGACTTCCAGAGAGTTCAGAGCAGAGCAGATGATGATACATAGGGGCTTCAAGAGG 2958
QY 1321 GAATTTGAAACTTAAGAACCTGTGAATCATGAGTACTCTTGAGACTGTGACAAATTTCTG 1380
DB 2959 GAATTTGAAACTTAAGAACCTGTGAATCATGAGTACTCTTGAGACTGTGACAAATTTCTG 3018
QY 1381 ACAGAGAGCCCTTTGGAAGGACTAGAGAACTCTACAGAGGCGCCAGAGAGCTGCTCCT 1440
DB 3019 ACAGAGAGCCCTTTGGAAGGACTAGAGAACTCTACAGAGGCGCCAGAGAGCTGCTCCT 3078
QY 1441 GAGGAGAGCCCGAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACT 1500
DB 3079 GAGGAGAGCCCGAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACT 3138
QY 1501 GAGTGGGAAAATTAAGAACTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCT 1560
DB 3139 GAGTGGGAAAATTAAGAACTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCT 3198
QY 1561 GAAAGACTCCAGGAACTTCAAGAGGCCACGAGTGAAGTGGACCTCAAGCTGCGCCAGCT 1620
DB 3199 GAAAGACTCCAGGAACTTCAAGAGGCCACGAGTGAAGTGGACCTCAAGCTGCGCCAGCT 3258
QY 1621 GAGGTGATCAAGGATCTGGCAGCCGCTGGCGGATCTCTCATTTGACTCTCTCCAGAT 1680
DB 3259 GAGGTGATCAAGGATCTGGCAGCCGCTGGCGGATCTCTCATTTGACTCTCTCCAGAT 3318
QY 1681 CACCTCGAGAAAGTCAAGCACTTCGAGAGAAATTTGGCCTCTGAAAGAGAACGTGAGC 1740
DB 3319 CACCTCGAGAAAGTCAAGCACTTCGAGAGAAATTTGGCCTCTGAAAGAGAACGTGAGC 3378
QY 1741 CACGTCAATGACTTGTGCGCAGCTTACCACTTTGGCATTGAGCTCTCCAGCTATAAC 1800
DB 3379 CACGTCAATGACTTGTGCGCAGCTTACCACTTTGGCATTGAGCTCTCCAGCTATAAC 3438
QY 1801 CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGAC 1860
DB 3439 CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGAC 3498
QY 1861 CGAGTCAGGAGCTGTGATGAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTCTT 1920
DB 3499 CGAGTCAGGAGCTGTGATGAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTCTT 3558
QY 1921 TCCAGCTCTGTCAGGCTCCCTGGGAGAGGCGCATCTCGCCAAACAAAGTGCCTTACT 1980
DB 3559 TCCAGCTCTGTCAGGCTCCCTGGGAGAGGCGCATCTCGCCAAACAAAGTGCCTTACT 3618
QY 1981 ATCAACACAGAGACTCAACA 2001
DB 3619 ATCAACACAGAGACTCAACA 3639
```

RESULT 7

```
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match      85.4%; Score 1709; DB 13; Length 5060;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conserved 0; Mismatches 0; Indels 141; Gaps 1;

QY 1 GGCAGCTTCATTGATGAGAGTGAAGTAAACCTGACCGGTATCAACAGCTTTTAGAGAA 60
DB 1992 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGACCGGTATCAACAGCTTTTAGAGAA 2051
QY 61 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTCGAGCAGACAGGAGATTTCTAAT 120
DB 2052 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTCGAGCAGACAGGAGATTTCTAAT 2111
QY 121 GATGTGGAAGTGGTGAAGACCCAGTTTCATCTACTCATGAGGGGTACATGATGATTTGACA 180
DB 2112 GATGTGGAAGTGGTGAAGACCCAGTTTCATCTACTCATGAGGGGTACATGATGATTTGACA 2171
QY 181 GCCCATCAGGCGCGGTTGGTAATATTCCTCAATTTGGGAAGTAACTGATTTGGACAGGA 240
DB 2172 GCCCATCAGGCGCGGTTGGTAATATTCCTCAATTTGGGAAGTAACTGATTTGGACAGGA 2231
QY 241 AAATTTATCAGAGATGGAAGAACTGCAAGTACAGAGCAGATGAATCTCTAAATTCAGA 300
DB 2232 AAATTTATCAGAGATGGAAGAACTGCAAGTACAGAGCAGATGAATCTCTAAATTCAGA 2291
QY 301 TGGGAATGCTTCAGGCTAGCTAGCATGGAAAAAACAAGCAATTTACATAGAGTTTAAATG 360
DB 2292 TGGGAATGCTTCAGGCTAGCTAGCATGGAAAAAACAAGCAATTTACATAGAGTTTAAATG 2351
QY 361 GATCTCCAGAAATCAGAACTGAAAGAGTTGATGACTGCTAAACAAAACAGAGAAAGA 420
DB 2352 GATCTCCAGAAATCAGAACTGAAAGAGTTGATGACTGCTAAACAAAACAGAGAAAGA 2411
QY 421 ACAAGGAAATGGAGAGAGGCTCTTGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 480
DB 2412 ACAAGGAAATGGAGAGAGGCTCTTGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 2471
QY 481 CAACAAATTAAGGTGCTTCAAGAGATCTAGACAGAGACAGTCAAGGTCAATTTCTCTC 540
DB 2472 CAACAAATTAAGGTGCTTCAAGAGATCTAGACAGAGACAGTCAAGGTCAATTTCTCTC 2531
QY 541 ACTCAATGCTGTGTAGTTCATGAATCTAGTGGAGATCACGCAACTGCTCTTTGGAA 600
DB 2532 ACTCAATGCTGTGTAGTTCATGAATCTAGTGGAGATCACGCAACTGCTCTTTGGAA 2591
QY 601 GAACAACTTAAGGTATTTGGAGATCGATGGGCAACATCTGTAGTGGACAGAGACCGC 660
DB 2592 GAACAACTTAAGGTATTTGGAGATCGATGGGCAACATCTGTAGTGGACAGAGACCGC 2651
QY 661 TGGGTCTTTTACAAGACCCAGCCTGACCTAGCTCTCTGGACTGACCACCTATTGGAGCCTCT 720
DB 2652 TGGGTCTTTTACAAGC----- 2669
QY 721 CCTACTCAGACTCTTACTCTGTTGTCACAACTGTGGTTACTAAGGAAACTGCCATCTCC 780
DB 2670 ----- 2669
QY 781 AAACAGAAATGCCATCTTCTTGTGTTGGAGGTACTACTCATAGATTACTGCAACAG 840
DB 2670 ----- 2690
QY 841 TTCCCTCTGGACCTGGAAAAAGTTTCTTGTGCTGCTTACAGAGCTGAAACACTGCCAAT 900
DB 2691 TTCCCTCTGGACCTGGAAAAAGTTTCTTGTGCTGCTTACAGAGCTGAAACACTGCCAAT 2750
QY 901 GTCCTACAGATGCTACCCGTAGAGAAAGCTCTCTAGAGACTCCCAAGGAGTAAAGAG 960
DB 2751 GTCCTACAGATGCTACCCGTAGAGAAAGCTCTCTAGAGACTCCCAAGGAGTAAAGAG 2810
```

```
961 CTGATGAAACAAATGGCAAGACCTCAAAGTGAAATTTGAAGCTCACACAGATGTTTATCAC 1020
2811 CTGATGAAACAAATGGCAAGACCTCAAAGTGAAATTTGAAGCTCACACAGATGTTTATCAC 2870
1021 AACCTGGATGAAACAGCCAAATAATCTCGATCCCTGGAAAGTTCGATGATGACGATC 1080
2871 AACCTGGATGAAACAGCCAAATAATCTCGATCCCTGGAAAGTTCGATGATGACGATC 2930
1081 CTGTTACAAAGAGCTTTGGATTAACATGAATCTCAAGTGAGTGAATCTCGGAAAGTCT 1140
2931 CTGTTACAAAGAGCTTTGGATTAACATGAATCTCAAGTGAGTGAATCTCGGAAAGTCT 2990
1141 CTCACATTAAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGACCTTTCT 1200
2991 CTCACATTAAGTCCCATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGACCTTTCT 3050
1201 CTCGAGGAATCTCTGGTGGCTTACAGCTGAAGATGATGAATTAAGCCGCGAGGACCT 1260
3051 CTCGAGGAATCTCTGGTGGCTTACAGCTGAAGATGATGAATTAAGCCGCGAGGACCT 3110
1261 ATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGTATACATAGGCGCTTCAAGAGG 1320
3111 ATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGATGTATACATAGGCGCTTCAAGAGG 3170
1321 GAATTCGAAACCTTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
3171 GAATTCGAAACCTTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 3230
1381 ACAGAGCAGCTTTGGAAGGACTTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCT 1440
3231 ACAGAGCAGCTTTGGAAGGACTTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCT 3290
1441 GAGGAGAGCCAGAAATGTCTACGCGCTTTCAGAAAGCAGGCTGAGGAGTCAATACT 1500
3291 GAGGAGAGCCAGAAATGTCTACGCGCTTTCAGAAAGCAGGCTGAGGAGTCAATACT 3350
1501 GAGTGGGAAATTTGAACCTGCACTCCGCTGAGTGCAGAGAAATAGATGAGACCTTT 1560
3351 GAGTGGGAAATTTGAACCTGCACTCCGCTGAGTGCAGAGAAATAGATGAGACCTTT 3410
1561 GAAAGACTCCAGAACTTCAAGAGGCCAGGATGAGCTTGACCTCAAGCTGCGCAAGCT 1620
3411 GAAAGACTCCAGAACTTCAAGAGGCCAGGATGAGCTTGACCTCAAGCTGCGCAAGCT 3470
1621 GAGGTATCAAGGGATCTTGGAGCCGCTGGGCGATCTCTCTCATTTGACTCTCTCAAGAT 1680
3471 GAGGTATCAAGGGATCTTGGAGCCGCTGGGCGATCTCTCTCATTTGACTCTCTCAAGAT 3530
1681 CACCTCGAGAAAGTCAAGGCACTTTCAGAGAGAAATTCGCGCTCTGAAAGAGAACTGAGC 1740
3531 CACCTCGAGAAAGTCAAGGCACTTTCAGAGAGAAATTCGCGCTCTGAAAGAGAACTGAGC 3590
1741 CAGGTCAATGACTTCTGCGCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCCTTAAC 1800
3591 CAGGTCAATGACTTCTGCGCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCCTTAAC 3650
1801 CTCAGCACTCTGAGAGCTTGAACACAGATGAAGCTTCTGAGGTGGCCGCTCCAGGAC 1860
3651 CTCAGCACTCTGAGAGCTTGAACACAGATGAAGCTTCTGAGGTGGCCGCTCCAGGAC 3710
1861 CGAGTCAGGAGCTGATGAAGCCACAGGACTTTTGGTCCAGCATCTCAGCACTTTCTT 1920
3711 CGAGTCAGGAGCTGATGAAGCCACAGGACTTTTGGTCCAGCATCTCAGCACTTTCTT 3770
1921 TCCAGCTCTGTCAGGCTCCCTGGAGAGGACCTCTCCGCAAAACAAAGTGCCTTACTAT 1980
3771 TCCAGCTCTGTCAGGCTCCCTGGAGAGGACCTCTCCGCAAAACAAAGTGCCTTACTAT 3830
1981 ATCAACCCAGAGACTCAAAACA 2001
3831 ATCAACCCAGAGACTCAAAACA 3851
```

```
RESULT 8
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE11142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match      83.1%; Score 1662.8; DB 13; Length 4182;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

Qy 1 GGCAGTTCATGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA 60
Db 1000 GGCAGTTCATGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA 1059
Qy 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGACGACCAAGGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGACGACCAAGGAGAGATTTCTAAT 1119
Qy 121 GATGTGAAGTGTGAAAGACACGATTTTCATCTCATGAGGGTACATGATGATTTTGACA 180
Db 1120 GATGTGAAGTGTGAAAGACACGATTTTCATCTCATGAGGGTACATGATGATTTTGACA 1179
Qy 181 GCCCATCAGGCGCGGTTGGTAAATTTCTACAAATTCGGAAGTAAAGCTGATTTGAAACAGGA 240
Db 1180 GCCCATCAGGCGCGGTTGGTAAATTTCTACAAATTCGGAAGTAAAGCTGATTTGAAACAGGA 1239
Qy 241 AAATATTCAGAAATGAAGAACTGAAGTACAAAGACAGATGAATCTCTTAATTCAGAA 300
Db 1240 AAATATTCAGAAATGAAGAACTGAAGTACAAAGACAGATGAATCTCTTAATTCAGAA 1299
Qy 301 TGGGAATCCCTCAGGCTAGCTAGCATGGAAACCAAGCAATTTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATCCCTCAGGCTAGCTAGCATGGAAACCAAGCAATTTTACATAGAGTTTAAATG 1359
Qy 361 GATCTCCAGAAATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAA 420
Db 1360 GATCTCCAGAAATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAAAACAGAAAGAA 1419
Qy 421 ACAAGGAAATGAGAGAGAGCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTA 480
Db 1420 ACAAGGAAATGAGAGAGAGCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTA 1479
Qy 481 CAACAACTAAGGTGCTTCAAGAGATCTAGAAACAAAGAACAAAGTCAAGGTCAATTTCTC 540
Db 1480 CAACAACTAAGGTGCTTCAAGAGATCTAGAAACAAAGAACAAAGTCAAGGTCAATTTCTC 1539
Qy 541 ACTCAGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 1540 ACTCAGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599
Qy 601 GAACAACTTAAGGTGATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCCG 660
Db 1600 GAACAACTTAAGGTGATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCCG 1659
Qy 661 TGGGTCTTTTACAGACACGCTGAC-----CTAGCTCTCTGACTGACACT 708
Db 1660 TGGGTCTTTTACAGACACATCTCTCAATGGCAACGCTTCTTACTGAAGAACAGTGCCTT 1719
```

QY 709 ATTGGAGCCTCTCTACTCAGACTGTTACTCTGGTGACACA-----ACCTGTGGTT 759
Db 1720 TTTAGTGCATGGCTTTTCAGAAAAGAGATGAGTGAGTACAGATTACACAACTGGCTTT 1779
QY 760 ACTAAGAACTGCCATCTC----- 779
Db 1780 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAGCTGGCGGTTTAAAGCGGATCTA 1839
QY 780 -----CAAACTAGAAATGCCATCTTCC 801
Db 1840 GAAAAGAAAAGCAATCTCATGGGCAAACTGTATTCACTCAAAAGATCTTTTCAACA 1899
QY 802 TTGATGTTGGAG----- 813
Db 1900 CTGAAGATAAGTCAGTGAGCCAGAGACGGAAGCATGGCTGGATAAATTTTCCCGGTGT 1959
QY 814 -----GTACCTTACTCATAGATTACTGCAA 837
Db 1960 TGGGATAATTTAGTCCAAAAAATTTGAAAAGAGTACAGCAACAGACTCATAGATTACTGCAA 2019
QY 838 CAGTTCCTCCCTGACCTGGAAAAGTTTCTGCTGGCTTACAGAAAGCTGAAACAACTGCC 897
Db 2020 CAGTTCCTCCCTGACCTGGAAAAGTTTCTGCTGGCTTACAGAAAGCTGAAACAACTGCC 2079
QY 898 AATGTCCTACAGGATGCTACCCGTAAAGGAAAGCTCCTTAGAAGACTCCAGGGAATAAA 957
Db 2080 AATGTCCTACAGGATGCTACCCGTAAAGGAAAGCTCCTTAGAAGACTCCAGGGAATAAA 2139
QY 958 GAGCTGATGAACAATGGGAAGACCTCCAAAGGTGAATTTGAAGCTCACACAGATGTTAT 1017
Db 2140 GAGCTGATGAACAATGGGAAGACCTCCAAAGGTGAATTTGAAGCTCACACAGATGTTAT 2199
QY 1018 CACAACCTGGATGAAACAGCCAAAATACTCTGAGATCCCTGGAAAGTTCCGATGATGA 1077
Db 2200 CACAACCTGGATGAAACAGCCAAAATACTCTGAGATCCCTGGAAAGTTCCGATGATGA 2259
QY 1078 GTCTGTACAAAGAGCTTTGGATACATGAATCAAGTGAGTGAACTTCGGAAAAAG 1137
Db 2260 GTCTGTACAAAGAGCTTTGGATACATGAATCAAGTGAGTGAACTTCGGAAAAAG 2319
QY 1138 TCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAAGCTCTGACCTT 1197
Db 2320 TCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACAGTGGAAAGCTCTGACCTT 2379
QY 1198 TCTCTCAGGAATCTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGCA 1257
Db 2380 TCTCTCAGGAATCTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGCA 2439
QY 1258 CCTATTGGAGCGGACTTTCAGCAGTTCCAGAGCAGAGCAGTGTACATAGGGCCTTCAAG 1317
Db 2440 CCTATTGGAGCGGACTTTCAGCAGTTCCAGAGCAGAGCAGTGTACATAGGGCCTTCAAG 2499
QY 1318 AGGGAATTGAAAACCTAAGAACCTGTAAATCATAGTACTCTTGAGACTGTACGAATATT 1377
Db 2500 AGGGAATTGAAAACCTAAGAACCTGTAAATCATAGTACTCTTGAGACTGTACGAATATT 2559
QY 1378 CTGACAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCT 1437
Db 2560 CTGACAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCT 2619
QY 1438 CCTGAGGAGAGCCAGAAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGAGGTCAAT 1497
Db 2620 CCTGAGGAGAGCCAGAAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGAGGTCAAT 2679
QY 1498 ACTGAGTGGAAAAATTTGAACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1557
Db 2680 ACTGAGTGGAAAAATTTGAACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2739
QY 1558 CTTGAAAGACTCCAGGAATCTTCAAGAGGCCACCGATGAGCTGAGCTCAAGCTCGGCCAA 1617
Db 2740 CTTGAAAGACTCCAGGAATCTTCAAGAGGCCACCGATGAGCTGAGCTCAAGCTCGGCCAA 2799

QY 1618 GCTCAGGTGATCAAGGATCTCTGGAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAA 1677
Db 2800 GCTCAGGTGATCAAGGATCTCTGGAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAA 2859
QY 1678 GATCAGCTCGAGAAATCTAAGCAGCTTCGAGGAGAAATTCGCGCTCTGAAGAGACGTG 1737
Db 2860 GATCAGCTCGAGAAATCTAAGCAGCTTCGAGGAGAAATTCGCGCTCTGAAGAGACGTG 2919
QY 1738 AGCCAGCTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTAT 1797
Db 2920 AGCCAGCTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTAT 2979
QY 1798 AACTCAGCAGCTCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGCGGTCGAG 1857
Db 2980 AACTCAGCAGCTCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAGGTGCGGTCGAG 3039
QY 1858 GACCGAGTCAGGAGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTT 1917
Db 3040 GACCGAGTCAGGAGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTT 3099
QY 1918 CTTTCCAGCTCTCTCAGGCTCCCTGGAGAGAGCCATCTCGCCAAAACAAAGTCCCTTAC 1977
Db 3100 CTTTCCAGCTCTCTCAGGCTCCCTGGAGAGAGCCATCTCGCCAAAACAAAGTCCCTTAC 3159
QY 1978 TATATCAACCCAGAGACTCAACA 2001
Db 3160 TATATCAACCCAGAGACTCAACA 3183

RESULT 9
US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27

Query Match 83.1%; Score 1662.8; DB 13; Length 5149;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

QY 1 GCGAGTTTCATGTATGAGAGTGAAGTAAACCTCGAGCTTATCAACAGCTTTTAGAAGAA 60
Db 1757 GCGAGTTTCATGTATGAGAGTGAAGTAAACCTCGAGCTTATCAACAGCTTTTAGAAGAA 1816
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGACACAAAGAGAGATTTCTTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGACACAAAGAGAGATTTCTTAAT 1876
QY 121 GATGTGAAAGTGGTGAAGACCCAGTTTTCATCTATCATGAGGGGTACATGATGGATTGACA 180
Db 1877 GATGTGAAAGTGGTGAAGACCCAGTTTTCATCTATCATGAGGGGTACATGATGGATTGACA 1936
QY 181 GCCCATCAGGCGCGGTTGGTAAATATTTCTACAATTTGGAAAGTAAAGCTGGAACAGGA 240
Db 1937 GCCCATCAGGCGCGGTTGGTAAATATTTCTACAATTTGGAAAGTAAAGCTGGAACAGGA 1996
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 300
Db 1997 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 2056

QY 301 TGGGAATGCTTCAGGGTAGCTAGCATGGAAAAAACAAGCAATTTACATAGATTTTAATG 360
Db 2057 TGGGAATGCTTCAGGGTAGCTAGCATGGAAAAAACAAGCAATTTACATAGATTTTAATG 2116
QY 361 GATCTCCAGATCAGAACTGAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGA 420
Db 2117 GATCTCCAGATCAGAACTGAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGA 2176
QY 421 ACAAGGAAATGAGGAGAGGCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTA 480
Db 2177 ACAAGGAAATGAGGAGAGGCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTA 2236
QY 481 CAACAACTAAGGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTTC 540
Db 2237 CAACAACTAAGGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTTC 2296
QY 541 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
Db 2297 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2356
QY 601 GAACAACTAAGGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTTC 660
Db 2357 GAACAACTAAGGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTTC 2416
QY 661 TGGGTCTTTTACAGACAGGCTGAC-----CTAGCTCTGACCTGACCACT 708
Db 2417 TGGGTCTTTTACAGACAGGCTGAC-----CTAGCTCTGACCTGACCACT 2476
QY 709 ATTGGAGCTCTCTCTACTCAGATGTTTACTCTGCTGACACA-----ACCTGGTGT 759
Db 2477 TTTAGTGCATGGCTTTTACAGAAAGAGATGCTAGTGAACAGATTCACACAACTGCTTT 2536
QY 760 ACTAAGGAATGCCATCTC----- 779
Db 2537 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTA 2596
QY 780 -----CAAAGTGAATGCCATCTTCC 801
Db 2597 GAAAGAAAAGCAATCCATGGGCAACTGTATTCTACTCAACAGATCTTCTTCAACA 2656
QY 802 TTGATGTTGGAG----- 813
Db 2657 CTGAAGAATAAGTCAGTGAACCCAGAGAGCAATGGCTGGATAACTTTGCCCGGTGT 2716
QY 814 -----GTACTACTCATAGATTACTGCAA 837
Db 2717 TGGGATTAATTTATGCCAAAACCTTGAAGAGATACAGCAGACTCATAGATTACTGCAA 2776
QY 838 CAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGGCTTACAGAACTGAAACAACTGCC 897
Db 2777 CAGTTCCTCCCTGGACCTGGAAAAAGTTTCTTGGCTTACAGAACTGAAACAACTGCC 2836
QY 898 AATGCTCTACAGATGCTACCGTGAAGAAAGCTCTAGAAAGCTTCAAGGAGTAAAA 957
Db 2837 AATGCTCTACAGATGCTACCGTGAAGAAAGCTTCTAGAAAGCTTCAAGGAGTAAAA 2896
QY 958 GAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGAATTCAGCTCACACAGATTTAT 1017
Db 2897 GAGCTGATGAAACAAATGGCAAGACCTCCAAGGTGAATTCAGCTCACACAGATTTAT 2956
QY 1018 CAACTCTGATGAAACAAAGCAAAATCCTGAGATCCCTGGAAGTTCGAGATGATCA 1077
Db 2957 CAACTCTGATGAAACAAAGCAAAATCCTGAGATCCCTGGAAGTTCGAGATGATCA 3016
QY 1078 GTCCGTGTTACAAAGACCTTTGGATACATCACTTCAAGTGGAGTCACTTGGAAAAAG 1137
Db 3017 GTCCGTGTTACAAAGACCTTTGGATACATCACTTCAAGTGGAGTGAATTCGAAAAAG 3076
QY 1138 TCTCTCAACATTTAGGTCCCATTTTGGAGCCAGTTCTGACAGTGGAGGCTCTGACCTT 1197
Db 3077 TCTCTCAACATTTAGGTCCCATTTTGGAGCCAGTTCTGACAGTGGAGGCTCTGACCTT 3136

QY 1198 TCTCTCAGGAACTTCTGGTCTGGCTACAGCTGAAAGATGATGAATTAAGCCGCAAGCA 1257
Db 3137 TCTCTCAGGAACTTCTGGTCTGGCTACAGCTGAAAGATGATGAATTAAGCCGCAAGCA 3196
QY 1258 CCTATTGAGGCGGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGAGGCTTCAAG 1317
Db 3197 CCTATTGAGGCGGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGAGGCTTCAAG 3256
QY 1318 AGGGAATTTGAAAACTTAAGAAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTT 1377
Db 3257 AGGGAATTTGAAAACTTAAGAAACCTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTT 3316
QY 1378 CTGACAGAGCAGCCTTTTGAAGAGACTAGAGAACTCTACAGAGGAGCCAGAGAGCTGCT 1437
Db 3317 CTGACAGAGCAGCCTTTTGAAGAGACTAGAGAACTCTACAGAGGAGCCAGAGAGCTGCT 3376
QY 1438 COTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAT 1497
Db 3377 COTGAGGAGAGAGCCAGAACTGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAT 3436
QY 1498 ACTGAGTGGGAAAAATTGAACCTGCACTCCGCTCACTGGGAGAGAAAAATAGATGAGACC 1557
Db 3437 ACTGAGTGGGAAAAATTGAACCTGCACTCCGCTCACTGGGAGAGAAAAATAGATGAGACC 3496
QY 1558 CTTGAAAGACCTCCAGGAGCTTCAAGAGGCCACGATGAGCTGGACCTCAAGCTCGCCAA 1617
Db 3497 CTTGAAAGACCTCCAGGAGCTTCAAGAGGCCACGATGAGCTGGACCTCAAGCTCGCCAA 3556
QY 1618 GCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGGCGATCTCTCTCATTTGACTCTCTCAA 1677
Db 3557 GCTGAGGTGATCAAGGGATCCTGGCAGCCGCTGGGCGATCTCTCTCATTTGACTCTCTCAA 3616
QY 1678 GATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGGCCTCTGAAAGAGAACTG 1737
Db 3617 GATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGGCCTCTGAAAGAGAACTG 3676
QY 1738 AGCCACGTCAATGACCTTGTCTGCCAGCTTACCCTTTGGGCACTCAGCTCTCACCGTAT 1797
Db 3677 AGCCACGTCAATGACCTTGTCTGCCAGCTTACCCTTTGGGCACTCAGCTCTCACCGTAT 3736
QY 1798 AACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTTGGAAGTGGCGCTGAG 1857
Db 3737 AACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTTGGAAGTGGCGCTGAG 3796
QY 1858 GACCGAGTCAGGAGCTGATGAAGCCACAGGCACTTTGTCCAGCATCTCAGCACTTT 1917
Db 3797 GACCGAGTCAGGAGCTGATGAAGCCACAGGCACTTTGTCCAGCATCTCAGCACTTT 3856
QY 1918 CTTTCCAGCTCTGTCCAGGCTCCTGGGAGAGCCATCTCGCCAAAACAAAGTGCCCTAC 1977
Db 3857 CTTTCCAGCTCTGTCCAGGCTCCTGGGAGAGCCATCTCGCCAAAACAAAGTGCCCTAC 3916
QY 1978 TATATCAACCAAGAGACTCAACA 2001
Db 3917 TATATCAACCAAGAGACTCAACA 3940

RESULT 10

US-10-149-736-41
; Sequence 41, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: US-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 5462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-41

Query Match 64.1%; Score 1283; DB 13; Length 5462;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1302; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Qy	691	GCTCCTGGACTGACCACTATTGAGGCTCTCTTACCTCAGCTGTTACTCTGTGTGACACAA	750
Db	1547	GCTCCTGGACTGACCACTATTGAGGCTCTCTTACCTCAGCTGTTACTCTGTGTGACACAA	1606
Qy	751	CTGTGTGTTACTAAGGAACTGCCATCTCCAACTAGAAATGCATCTTCCCTTGTGTTG	810
Db	1607	CCTGTGTGTTACTAAGGAACTGCCATCTCCAACTAGAAATGCATCTTCCCTTGTGTTG	1666
Qy	811	GAGTACCTACTCATGATTAATGCAACAGTTCCTCCCTGGACCTGGAAAGTTTCTTGCC	870
Db	1667	GAG-----CATGATTAATGCAACAGTTCCTCCCTGGACCTGGAAAGTTTCTTGCC	1717
Qy	871	TGGCTTACAGAGCTGAAACAACTGCCATGTCTTACAGATGCTACCCGTGAAGAAAG	930
Db	1718	TGGCTTACAGAGCTGAAACAACTGCCATGTCTTACAGATGCTACCCGTGAAGAAAG	1777
Qy	931	CTCCTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGCTTCCAAGT	990
Db	1778	CTCCTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGCTTCCAAGT	1837
Qy	991	GAAATTGAAGCTCACAGAGTGTATACAACTGGATGAAACAGCCGCAAAATCCTG	1050
Db	1838	GAAATTGAAGCTCACAGAGTGTATACAACTGGATGAAACAGCCGCAAAATCCTG	1897
Qy	1051	AGATCCCTGGAAGTTCGAGTGTATGAGTCTCTTACAAAGAGTGTGATTAACATGAAC	1110
Db	1898	AGATCCCTGGAAGTTCGAGTGTATGAGTCTCTTACAAAGAGTGTGATTAACATGAAC	1957
Qy	1111	TTCAAGTGGAGTGAACCTTCGAAAGAGTCTCTCAATAGTCCCATTTGNAAGCCAGT	1170
Db	1958	TTCAAGTGGAGTGAACCTTCGAAAGAGTCTCTCAATAGTCCCATTTGNAAGCCAGT	2017
Qy	1171	TCTGACCAAGTGGAGGCTCTGCACTTTCTCTCAGCAACTTCTGTGTGGCTTACAGTG	1230
Db	2018	TCTGACCAAGTGGAGGCTCTGCACTTTCTCTCAGCAACTTCTGTGTGGCTTACAGTG	2077
Qy	1231	AAAGATGATGAATTAAGCCGGCAGGCACTTATTTGGAGGAGCTTCCAGAGTTCAGAG	1290
Db	2078	AAAGATGATGAATTAAGCCGGCAGGCACTTATTTGGAGGAGCTTCCAGAGTTCAGAG	2137
Qy	1291	CAGAACGATGTACATAGGCTTCAAGAGGGAATTAAGAACTTAAGAACTGTATCATG	1350
Db	2138	CAGAACGATGTACATAGGCTTCAAGAGGGAATTAAGAACTTAAGAACTGTATCATG	2197
Qy	1351	AGTACTCTTGAGCTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGACTTAGAGAA	1410
Db	2198	AGTACTCTTGAGCTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGACTTAGAGAA	2257
Qy	1411	CTCTACAGGAGCCAGAGGCTGCTCTCTGAGAGAGAGCCAGAAATGTCTCTGGCTT	1470
Db	2258	CTCTACAGGAGCCAGAGGCTGCTCTCTGAGAGAGAGCCAGAAATGTCTCTGGCTT	2317
Qy	1471	CTACGAAGAGGCTGAGGAGGCTCAATCTAGTGGGAAATTTGAACCTGCACTCCGCT	1530
Db	2318	CTACGAAGAGGCTGAGGAGGCTCAATCTAGTGGGAAATTTGAACCTGCACTCCGCT	2377
Qy	1531	GACTGGCAGAGAAATATAGTGAACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCAG	1590
Db	2378	GACTGGCAGAGAAATATAGTGAACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCAG	2437

Qy	1591	GATGAGCTGAGACTCAAGCTGCGCCAAAGCTGAGTGATCAAGGATCTCTGGCAGCCGCTG	1650
Db	2438	GATGAGCTGAGACTCAAGCTGCGCCAAAGCTGAGTGATCAAGGATCTCTGGCAGCCGCTG	2497
Qy	1651	GGCGATCTCTCATTTGACTCTCTCCAGATCACTCGAGAAAGTCAAGGCACTTCGAGGA	1710
Db	2498	GGCGATCTCTCATTTGACTCTCTCCAGATCACTCGAGAAAGTCAAGGCACTTCGAGGA	2557
Qy	1711	GAAATTGGCTCTGAAAAGAGAGCTGAGCCACGTCATGACCTTGTCTGCCAGCTTACC	1770
Db	2558	GAAATTGGCTCTGAAAAGAGAGCTGAGCCACGTCATGACCTTGTCTGCCAGCTTACC	2617
Qy	1771	ACTTTGGCATTTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAAACACAGA	1830
Db	2618	ACTTTGGCATTTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAAACACAGA	2677
Qy	1831	TGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGG	1890
Db	2678	TGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAGGCAGCTGATGAAGCCACAGG	2737
Qy	1891	GACTTTGGTCCAGCATCTCAGCACTTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAGA	1950
Db	2738	GACTTTGGTCCAGCATCTCAGCACTTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAGA	2797
Qy	1951	GCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA	2001
Db	2798	GCCATCTCGCCAAACAAAGTGCCTACTATATCAACACGAGACTCAAAACA	2848

RESULT 11
US-10-149-736-42
; Sequence 42, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UN-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-42

Query Match 59.18; Score 1182.6; DB 13; Length 8689;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	813	GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCTG	872
Db	2992	GGAAGAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGCTG	3051
Qy	873	GCTTACAGAGCTGAAACAACTGCCATGTCTTACAGGATGCTACCGTGAAGGAAGGCT	932
Db	3052	GCTTACAGAGCTGAAACAACTGCCATGTCTTACAGGATGCTACCGTGAAGGAAGGCT	3111
Qy	933	CCTAGAAGACTTCCAAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGCTTCCAAGGTGA	992
Db	3112	CCTAGAAGACTTCCAAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGCTTCCAAGGTGA	3171
Qy	993	AATTGAAGCTCACAGATGTTTATCAACCTGGATGAAACAGCCAAATAATTCCTGAG	1052
Db	3172	AATTGAAGCTCACAGATGTTTATCAACCTGGATGAAACAGCCAAATAATTCCTGAG	3231

QY 1653 CGATCTCCTCATTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGGGAGA 1712
Db 8992 CGATCTCCTCATTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGGGAGA 8951
QY 1713 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTTGCTCGCCAGCTTACAC 1772
Db 8952 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTTGCTCGCCAGCTTACAC 9011
QY 1773 TTTGGGCAATTCAGCTCTCAACGCTATACCTCAGACCTCTGGAGACCTGAAACACGATG 1832
Db 9012 TTTGGGCAATTCAGCTCTCAACGCTATACCTCAGACCTCTGGAGACCTGAAACACGATG 9071
QY 1833 GAAGCTTCTGAGGTGCGCCGTGAGGACCGAGTCAGGACCTCATGAAAGCCACAGGGA 1892
Db 9072 GAAGCTTCTGAGGTGCGCCGTGAGGACCGAGTCAGGACCTCATGAAAGCCACAGGGA 9131
QY 1893 CTTTGGTCCAGACTCTCAGACATTTCTTCCAGTCTGTCCAGGTCCTCGGAGAGAGC 1952
Db 9132 CTTTGGTCCAGACTCTCAGACATTTCTTCCAGTCTGTCCAGGTCCTCGGAGAGAGC 9191
QY 1953 CATCTGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 2001
Db 9192 CATCTGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 9240

RESULT 13
US-10-149-736-44
; Sequence 44, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; CURRENT APPLICATION NUMBER: US/10/149,736
; FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 11443
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-44

Query Match 59.1%; Score 1182.6; DB 13; Length 11443;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACTCTCATAGATTACTGCAACAGTTCCTCCCTCGAAGTCTGAAAGTTCTTTCGCTG 872
Db 5746 GGAAGAACTCATAGATTACTGCAACAGTTCCTCCCTCGAAGTCTGAAAGTTCTTTCGCTG 5805
QY 873 GCTTACGAAGCTGAAACAACTGCGATGCTTCTACAGATGCTACCGTAAAGGAGGCT 932
Db 5806 GCTTACGAAGCTGAAACAACTGCGATGCTTCTACAGATGCTACCGTAAAGGAGGCT 5865
QY 933 CTTAGAACTCCAGGAGTAAAGAGCTGATGAAACAACTGGAAGACCTCCAAAGTGA 992
Db 5866 CTTAGAACTCCAGGAGTAAAGAGCTGATGAAACAACTGGAAGACCTCCAAAGTGA 5925
QY 993 AATTGAAGCTCACACAGATGTTTATCAACCTGGATGAAACAAAGCCAAATAATCTTGAG 1052
Db 5926 AATTGAAGCTCACACAGATGTTTATCAACCTGGATGAAACAAAGCCAAATAATCTTGAG 5985
QY 1053 ATCCCTGGAAGGTTCCGATGATGAGTCTGTTTCAAGAGGTTTGGATACATGACTT 1112

Db 5986 ATCCCTGGAAGGTTCCGATGATGAGTCTGTTTACAAAGAGCTTTGGATTAACATGACTT 6045
QY 1113 CAAGTGGAGTGAATTCGGAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTC 1172
Db 6046 CAAGTGGAGTGAATTCGGAAAGTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTC 6105
QY 1173 TGACCAAGTGAAGAGCTGTGCACCTTTCTTCCAGGAACTTCTGGTGGGTACAGCTGAA 1232
Db 6106 TGACCAAGTGAAGAGCTGTGCACCTTTCTTCCAGGAACTTCTGGTGGGTACAGCTGAA 6165
QY 1233 AGATGATGATTAAGCCGAGGACCTATTGAGGCGACTTTTCCAGCAGTTCAGAGCA 1292
Db 6166 AGATGATGATTAAGCCGAGGACCTATTGAGGCGACTTTTCCAGCAGTTCAGAGCA 6225
QY 1293 GAAAGATGTAATAGGAGGCTTCAAGAGGGAATTTGAAAGCTTAAAGAACTTCAATCATGAG 1352
Db 6226 GAAAGATGTAATAGGAGGCTTCAAGAGGGAATTTGAAAGCTTAAAGAACTTCAATCATGAG 6285
QY 1353 TACTCTTGAAGTCTGATGATTAATTTCTGACAGAGCAGCTTTTGGAGGACTAGAGAACT 1412
Db 6286 TACTCTTGAAGTCTGATGATTAATTTCTGACAGAGCAGCTTTTGGAGGACTAGAGAACT 6345
QY 1413 CTACAGAGGCCAGAGAGTCTCTTCTGAGGAGAGAGCCAGAACTTCACTCGGCTTCT 1472
Db 6346 CTACAGAGGCCAGAGAGTCTCTTCTGAGGAGAGAGCCAGAACTTCACTCGGCTTCT 6405
QY 1473 ACGAAAGCAGCTGAGAGGTCATATCTGAGTGGGAAAAATTTGAACTCGCTCGCTG 1532
Db 6406 ACGAAAGCAGCTGAGAGGTCATATCTGAGTGGGAAAAATTTGAACTCGCTCGCTG 6465
QY 1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGGCCAGGGA 1592
Db 6466 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGGCCAGGGA 6525
QY 1593 TGAGTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGATTCCTGGCAGCCCTGGG 1652
Db 6526 TGAGTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGATTCCTGGCAGCCCTGGG 6585
QY 1653 CGATCTCCTCATTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 1712
Db 6586 CGATCTCCTCATTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGA 6645
QY 1713 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTCTCGCCAGCTTACAC 1772
Db 6646 AATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTCTCGCCAGCTTACAC 6705
QY 1773 TTTGGGCAATTCAGCTCTCACCGTATACCTCAGCAGCTCTGGAAGACCTGAACACGATG 1832
Db 6706 TTTGGGCAATTCAGCTCTCACCGTATACCTCAGCAGCTCTGGAAGACCTGAACACGATG 6765
QY 1833 GAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 1892
Db 6766 GAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCAGGCAGCTGCATGAAGCCACAGGGA 6825
QY 1893 CTTTGGTCCAGACTCTCAGCAGCTTTCTTTCACGCTGTGTCCAGGTCCTCGGAGAGAGC 1952
Db 6826 CTTTGGTCCAGACTCTCAGCAGCTTTCTTTCACGCTGTGTCCAGGTCCTCGGAGAGAGC 6885
QY 1953 CATCTGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 2001
Db 6886 CATCTGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 6934

RESULT 14
US-10-149-736-47
; Sequence 47, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 11443
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-47

; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 12057
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-47

Query Match 59.1%; Score 1182.6; DB 13; Length 12057;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCGCCCTGGACCTGGAAAAGTTCTTGCCCTG 872
DB 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCGCCCTGGACCTGGAAAAGTTCTTGCCCTG 8319
QY 873 GCTTACAGAGCTGAACAACTGCTACAGAGTCTACCCCTGAGGAGTCTGCAAGGAGCT 932
DB 8320 GCTTACAGAGCTGAACAACTGCTACAGAGTCTACCCCTGAGGAGTCTGCAAGGAGCT 8379
QY 933 CTTAGAGACTCCAGGAGTAAAGAGCTGATGAACATGGCAAGCTCCAGGTGA 992
DB 8380 CTTAGAGACTCCAGGAGTAAAGAGCTGATGAACATGGCAAGCTCCAGGTGA 8439
QY 993 AATTGAAGCTCACAGAGTGTATCAACCTGGTGAACAGCCAAATAATCTCTGAG 1052
DB 8440 AATTGAAGCTCACAGAGTGTATCAACCTGGTGAACAGCCAAATAATCTCTGAG 8499
QY 1053 ATCCCTGGAAGGTTCCGATGATGAGTCTCTTCAAGAGCTTTGATTAACATGAAT 1112
DB 8500 ATCCCTGGAAGGTTCCGATGATGAGTCTCTTCAAGAGCTTTGATTAACATGAAT 8559
QY 1113 CAAGTGGAGTGAACCTCGGAAAAGTCTCAACATGATGCTCCATTTGGAAGCAGTTC 1172
DB 8560 CAAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATGATGCTCCATTTGGAAGCAGTTC 8619
QY 1173 TGACCAAGTGAAGGCTGTGACCTTTCTCTCAGGAACTCTGCTGCTGCTGCTGCTGAA 1232
DB 8620 TGACCAAGTGAAGGCTGTGACCTTTCTCTCAGGAACTCTGCTGCTGCTGCTGCTGAA 8679
QY 1233 AGATGATGAATTAAGCCGGAGGACCTATTGAGGAGGACCTTCCAGAGTTCAGAGCA 1292
DB 8680 AGATGATGAATTAAGCCGGAGGACCTATTGAGGAGGACCTTCCAGAGTTCAGAGCA 8739
QY 1293 GAACGATGTACATAGGGCTTTCAAGAGGGAATTAAGAACTTAAGAACTTAATCATGAG 1352
DB 8740 GAACGATGTACATAGGGCTTTCAAGAGGGAATTAAGAACTTAAGAACTTAATCATGAG 8799
QY 1353 TACTCTTGAGAGTGPACGAATATTCTGACAGAGAGCTTTGGAAGGACTTAGAGAACT 1412
DB 8800 TACTCTTGAGAGTGPACGAATATTCTGACAGAGAGCTTTGGAAGGACTTAGAGAACT 8859
QY 1413 CTACCAGAGCCAGAGAGTGCCTCTGAGGAGAGGAGCCAGAACTGCTACCTGGCTTCT 1472
DB 8860 CTACCAGAGCCAGAGAGTGCCTCTGAGGAGAGGAGCCAGAACTGCTACCTGGCTTCT 8919
QY 1473 ACGAAAGCAGGCTGAGGAGGTCAATATCTAGTGGGAAAATTTGAACCTGCATCCGCTGA 1532
DB 8920 ACGAAAGCAGGCTGAGGAGGTCAATATCTAGTGGGAAAATTTGAACCTGCATCCGCTGA 8979
QY 1533 CTGGCAGAGAAAATAGATGAGACCTTGAAGCTCCAGAACTTCAAGAGGCCAGGA 1592
DB 8980 CTGGCAGAGAAAATAGATGAGACCTTGAAGCTCCAGAACTTCAAGAGGCCAGGA 9039
QY 1593 TGAGCTGGACCTCAAGCTGCGCAGAGTGAAGTGAATCAAGGAGTCTTGGGAGCCGCTGGG 1652

DB 9040 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCTTGGCAGCCCGTGG 9099
QY 1653 CGATCTCTCTCATTTGACTCTCTCCAAAGATCACTCGAGAAAGTCAAGGACATTCGAGGAGA 1712
DB 9100 CGATCTCTCTCATTTGACTCTCTCCAAAGATCACTCGAGAAAGTCAAGGACATTCGAGGAGA 9159
QY 1713 AATTGGGCTCTGAAAGAGAGCTGAGGACAGTGAATGACCTTGTCTGCGCAGCTTACCAC 1772
DB 9160 AATTGGGCTCTGAAAGAGAGCTGAGGACAGTGAATGACCTTGTCTGCGCAGCTTACCAC 9219
QY 1773 TTTGGGCATTGAGCTCTCACCGTATAAAGCTCAGCACTCTGGAAGACCTGAAACACAGATG 1832
DB 9220 TTTGGGCATTGAGCTCTCACCGTATAAAGCTCAGCACTCTGGAAGACCTGAAACACAGATG 9279
QY 1833 GAAGCTTTGCGAGGTGGCGCTGCGAGGACCGAGTCAAGCAGCTGCAATGAAGCCCAAGGA 1892
DB 9280 GAAGCTTTGCGAGGTGGCGCTGCGAGGACCGAGTCAAGCAGCTGCAATGAAGCCCAAGGA 9339
QY 1893 CTTTGGTCCAGCATCTCAGCACTTTTCCACGCTCTGTCAGGGTCCCTGGGAGAGAGC 1952
DB 9340 CTTTGGTCCAGCATCTCAGCACTTTTCCACGCTCTGTCAGGGTCCCTGGGAGAGAGC 9399
QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 2001
DB 9400 CATCTCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACA 9448
RESULT 15
US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-22
Query Match 59.1%; Score 1182.6; DB 10; Length 13957;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCGCCCTGGACCTGGAAAAGTTCTTGCCCTG 872
DB 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCGCCCTGGACCTGGAAAAGTTCTTGCCCTG 8319
QY 873 GCTTACAGAGCTGAACAACTGCTACAGAGTCTACCCCTGAGGAGTCTTCAAGGAGCT 932
DB 8320 GCTTACAGAGCTGAACAACTGCTACAGAGTCTACCCCTGAGGAGTCTTCAAGGAGCT 8379
QY 933 CTTAGAGACTCCAGGAGTAAAGAGCTGATGAACATGGCAAGCTCCAGGTGA 992
DB 8380 CTTAGAGACTCCAGGAGTAAAGAGCTGATGAACATGGCAAGCTCCAGGTGA 8439
QY 993 AATTGAAGCTCACAGAGTGTATCAACCTGGTGAACAGCCAAATAATCTCTGAG 1052
DB 8440 AATTGAAGCTCACAGAGTGTATCAACCTGGTGAACAGCCAAATAATCTCTGAG 8499
QY 1053 ATCCCTGGAAGGTTCCGATGATGAGTCTCTTCAAGAGCTTTGATTAACATGAAT 1112
DB 8500 ATCCCTGGAAGGTTCCGATGATGAGTCTCTTCAAGAGCTTTGATTAACATGAAT 8559

Qy	1113	CAAGTGAGTGAACCTTCGGAAAGAGTCTCTCAACATTAGTCCCTCCATTTGGAAGCCAGTTTC	1172
Db	8560	CAAGTGAGTGAACCTTCGGAAAGAGTCTCTCAACATTAGTCCCTCCATTTGGAAGCCAGTTTC	8619
Qy	1173	TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAAGAACTTCTGGTGTGCTACAGCTGAA	1232
Db	8620	TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAAGAACTTCTGGTGTGCTACAGCTGAA	8679
Qy	1233	AGATGATGAATTAAAGCGGAGGACCTATTTGAGGCGACTTTCAGCAGCTTCAGAGCA	1292
Db	8680	AGATGATGAATTAAAGCGGAGGACCTATTTGAGGCGACTTTCAGCAGCTTCAGAGCA	8739
Qy	1293	GAACGATGTACATAGGCGCTTCAAGAGGGAATTTAAAGAACTTGTAAATCATGAG	1352
Db	8740	GAACGATGTACATAGGCGCTTCAAGAGGGAATTTAAAGAACTTGTAAATCATGAG	8799
Qy	1353	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACT	1412
Db	8800	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACT	8859
Qy	1413	CTACCAAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT	1472
Db	8860	CTACCAAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCT	8919
Qy	1473	ACGAAGCAGGCTGAGAGGTCATTAATCTGAGTGGGAAATTTGAACCTGCACTCCGCTGA	1532
Db	8920	ACGAAGCAGGCTGAGAGGTCATTAATCTGAGTGGGAAATTTGAACCTGCACTCCGCTGA	8979
Qy	1533	CTGCGAGAGAAATAGATGAGAGCCCTTGAAGAACTTCAAGAGCCAGCGA	1592
Db	8980	CTGCGAGAGAAATAGATGAGAGCCCTTGAAGAACTTCAAGAGCCAGCGA	9039
Qy	1593	TGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGTATCAAGGGATCTTGGCAGCCCGTGGG	1652
Db	9040	TGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGTATCAAGGGATCTTGGCAGCCCGTGGG	9099
Qy	1653	CGATCTCTCTATTTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGGA	1712
Db	9100	CGATCTCTCTATTTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGGA	9159
Qy	1713	AAATGCGCCTCTGAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCGCAGCTTACCAC	1772
Db	9160	AAATGCGCCTCTGAAGAGAACGTGAGCCACGTCAATGACCTTGTCTGCGCAGCTTACCAC	9219
Qy	1773	TTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTGAAACCCAGATG	1832
Db	9220	TTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTGAAACCCAGATG	9279
Qy	1833	GAAGCTTCTGCGAGTGGCGCTGAGAGCCGAGTCAGGAGCTGCATGAAGCCACAGGGA	1892
Db	9280	GAAGCTTCTGCGAGTGGCGCTGAGAGCCGAGTCAGGAGCTGCATGAAGCCACAGGGA	9339
Qy	1893	CTTTGTCCAGCATCTCAGACATTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGC	1952
Db	9340	CTTTGTCCAGCATCTCAGACATTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGC	9399
Qy	1953	CATCTGCCAAACAAAGTGCCTTACTATATCAACCCAGAGACTCAACA	2001
Db	9400	CATCTGCCAAACAAAGTGCCTTACTATATCAACCCAGAGACTCAACA	9448

Search completed: February 2, 2004, 17:38:32
Job time : 668.123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:11:40 ; Search time 123.697 Seconds
(without alignments)
7140.092 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000

Perfect score: 2001

Sequence: 1 ggcagttcattgatggagag.....tcaaccacagagactcaaca 2001

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Issued Patents NA.*
- 2: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182.6	59.1	5952	US-09-687-875A-1	Sequence 1, Appl
2	1182.6	59.1	13977	US-09-484-970B-60	Sequence 60, Appl
3	1004	50.2	19307	US-08-836-022A-10	Sequence 10, Appl
4	1004	50.2	19307	US-09-427-048A-10	Sequence 10, Appl
5	397	19.8	6045	US-09-091-501B-7	Sequence 7, Appl
6	397	19.8	10320	US-09-091-501B-9	Sequence 9, Appl
7	79.4	4.0	200	US-09-091-501B-5	Sequence 5, Appl
8	78.6	3.9	200	US-09-091-501B-4	Sequence 4, Appl
9	78.6	3.9	200	US-09-091-501B-6	Sequence 6, Appl
10	76.6	3.8	7218	US-08-232-463-14	Sequence 14, Appl
11	63.6	3.2	238	US-09-687-875A-13	Sequence 13, Appl
12	44	2.2	1230025	US-09-198-452A-1	Sequence 1, Appl
13	43.4	2.2	1175	US-09-107-312A-1186	Sequence 1186, Ap
14	42.8	2.1	1690	US-09-620-312D-69	Sequence 69, Appl
15	42.8	2.1	7812	US-09-368-590-1	Sequence 1, Appl
16	40.4	2.0	2223	US-08-257-073-4	Sequence 4, Appl
17	39.2	2.0	16995	US-08-961-527-82	Sequence 82, Appl
18	38.6	1.9	1751	US-09-620-312D-847	Sequence 847, App
19	38.6	1.9	1995	US-08-425-069-3	Sequence 3, Appl
20	38.6	1.9	1995	US-08-317-844B-3	Sequence 3, Appl
21	38.4	1.9	7672	US-08-220-132-24	Sequence 24, Appl
22	38.2	1.9	1131	5180810-3	Patent No. 5180810
23	38.2	1.9	1784	5180810-2	Patent No. 5180810
24	38	1.9	1394	US-09-247-155-76	Sequence 76, Appl
25	36.8	1.8	1886	5210183-1	Patent No. 5210183
26	36.6	1.8	1845	US-08-887-534A-22	Sequence 22, Appl
27	36.6	1.8	1845	US-09-527-431-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1

US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 654786
; GENERAL INFORMATION:
; APPLICANT: Liao, Paul
; APPLICANT: Liao, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPLICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match 59.1%; Score 1182.6; DB 4; Length 5952;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACTACTCATAGATTACTGCAACAGTTCCTCCCTGGAGCTGGAAGTTCTTGTCTG 872
DB 2946 GGAAGAACTCATAGATTACTGCAACAGTTCCTCCCTGGAGCTGGAAGTTCTTGTCTG 3005
QY 873 GCTTACAGAGCTGAAACAGTTCCTTACAGAGCTGCTTACAGAGCTGCTTACAGAGCT 932
DB 3006 GCTTACAGAGCTGAAACAGTTCCTTACAGAGCTGCTTACAGAGCTGCTTACAGAGCT 3065
QY 933 CTTAGAGCTCCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTTCCAAAGTGA 992
DB 3066 CTTAGAGCTCCCAAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTTCCAAAGTGA 3125
QY 993 AATTGAGCTCACAGAGTGTATATCAACCTGGATGAAACAGCCAAATATCTTGTAG 1052
DB 3126 AATTGAGCTCACAGAGTGTATATCAACCTGGATGAAACAGCCAAATATCTTGTAG 3185
QY 1053 ATCCCTGGAAGGTTCCGATGATGAGCTCTCTTCAAGAGCTTTGGATACATGAACCTT 1112

Db 3186 ATCCCTGGAAGGTTCCGATGATGAGTCTCTGTTCAAAAGACGTTTGGATAACATGAACCTT 3245
QY 1113 CAAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAAGTCCCATTTTGGAGCCAGTTC 1172
Db 3246 CAAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAAGTCCCATTTTGGAGCCAGTTC 3305
QY 1173 TGACCAAGTGAAGCGCTCTGCACCTTTCTCTGAGGAACTTTCTGGTGTGGCTTACAGCTGAA 1232
Db 3306 TGACCAAGTGAAGCGCTCTGCACCTTTCTCTGAGGAACTTTCTGGTGTGGCTTACAGCTGAA 3365
QY 1233 AGATGATGAATTAAGCCGCGAGGCACTTATGGAGGCACTTTCCAGCAGTTCAGAGCA 1292
Db 3366 AGATGATGAATTAAGCCGCGAGGCACTTATGGAGGCACTTTCCAGCAGTTCAGAGCA 3425
QY 1293 GAACGATGATACATAGGCGCTTCAAGAGGAACTTGAAGAACTTGAAGAACTTGAAGAACT 1352
Db 3426 GAACGATGATACATAGGCGCTTCAAGAGGAACTTGAAGAACTTGAAGAACTTGAAGAACT 3485
QY 1353 TACTCTTGAGCTGTACGAATATTTCTGACAGAGCAGGCTTTGGAGGACTAGAGAACT 1412
Db 3486 TACTCTTGAGCTGTACGAATATTTCTGACAGAGCAGGCTTTGGAGGACTAGAGAACT 3545
QY 1413 CTACAGGAGCCAGAGAGCTGCTCTGAGGAGAGAGCCAGAGATGTCACCTCGCTTCT 1472
Db 3546 CTACAGGAGCCAGAGAGCTGCTCTGAGGAGAGAGCCAGAGATGTCACCTCGCTTCT 3605
QY 1473 ACGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAAATTTGAACCTGCACTCGCTGTA 1532
Db 3606 ACGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAAATTTGAACCTGCACTCGCTGTA 3665
QY 1533 CTGGCAGAGAAAATAGATGAGACCTTTGAAGACTCCAGAGCTTCAAGAGCCAGCGA 1592
Db 3666 CTGGCAGAGAAAATAGATGAGACCTTTGAAGACTCCAGAGCTTCAAGAGCCAGCGA 3725
QY 1593 TGAGCTGGAACCTCAAGCTCGGCAAGCTGAGTGTATCAAGGATCCTGCGAGCCCGTGGG 1652
Db 3726 TGAGCTGGAACCTCAAGCTCGGCAAGCTGAGTGTATCAAGGATCCTGCGAGCCCGTGGG 3785
QY 1653 CGATCTCTCATGACTCTCTCAAGATCACTCGAGAAAGTCAAGGACTTCGAGGAGA 1712
Db 3786 CGATCTCTCATGACTCTCTCAAGATCACTCGAGAAAGTCAAGGACTTCGAGGAGA 3845
QY 1713 AATTGGCGCTCTGAAAGAGAACTGAGGCACTCAATGACCTTGTCTGCGCAGCTTACCAC 1772
Db 3846 AATTGGCGCTCTGAAAGAGAACTGAGGCACTCAATGACCTTGTCTGCGCAGCTTACCAC 3905
QY 1773 TTTGGGCAATCAGCTCTACCGTATTAACCTCAGCACTCTGGAAGACCTTGAACACCCAGATG 1832
Db 3906 TTTGGGCAATCAGCTCTACCGTATTAACCTCAGCACTCTGGAAGACCTTGAACACCCAGATG 3965
QY 1833 GAAGCTTCTGAGGTGGCGTGGAGGACGAGTCAAGGAGCTGATGAGCCAGGCAAGGCA 1892
Db 3966 GAAGCTTCTGAGGTGGCGTGGAGGACGAGTCAAGGAGCTGATGAGCCAGGCAAGGCA 4025
QY 1893 CTTTGGTCCAGCACTCAGCACTTTCTTCCAGCTGTCTCCAGGTCCTGCGAGGAGAGC 1952
Db 4026 CTTTGGTCCAGCACTCAGCACTTTCTTCCAGCTGTCTCCAGGTCCTGCGAGGAGAGC 4085
QY 1953 CATCTGCGCAAAACAAAGTGGCTTACTATATCAACCAAGGACTCAACA 2001
Db 4086 CATCTGCGCAAAACAAAGTGGCTTACTATATCAACCAAGGACTCAACA 4134

RESULT 2

US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US

; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match 59.1%; Score 1182.6; DB 4; Length 13977;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAGTTTCTTGCCTG 872
Db 8260 GGAAGAAACTCATAGATTACTGCAACAGTTCCCTCGACCTGGAAAGTTTCTTGCCTG 8319
QY 873 GCTTACAGAGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAAGGAAAGCT 932
Db 8320 GCTTACAGAGCTGAAACAACTGCCAATGTCTACAGGATGCTACCCGTAAAGGAAAGCT 8379
QY 933 CCTAGAGACTCCCAAGGAGTAAAGAGCTGTATGAACAACTCGCAAGACCTCCAGGTGA 992
Db 9380 CCTAGAGACTCCCAAGGAGTAAAGAGCTGTATGAACAACTCGCAAGACCTCCAGGTGA 8439
QY 993 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAACACGCAAAATCTCTGAG 1052
Db 8440 AATTGAAGCTCACACAGATGTTTATCACAACTGGATGAACACGCAAAATCTCTGAG 8499
QY 1053 ATCCCTGGAAGGTTCCGATGATGAGTCTCTGTATCAAGACGCTTTGGATTAACATGAACTT 1112
Db 8500 ATCCCTGGAAGGTTCCGATGATGAGTCTCTGTATCAAGACGCTTTGGATTAACATGAACTT 8559
QY 1113 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCCAGTTC 1172
Db 8560 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCCAGTTC 8619
QY 1173 TGACCAAGTGAAGCGCTGACACCTTTCTGAGGAACTTCTGTGTGGCTACAGCTGAA 1232
Db 8620 TGACCAAGTGAAGCGCTGACACCTTTCTGAGGAACTTCTGTGTGGCTACAGCTGAA 8679
QY 1233 AGATGATGAATTAAGCCGCGAGGCACTTATGGAGGCACTTTCCAGCAGTTCAGAGCA 1292
Db 8680 AGATGATGAATTAAGCCGCGAGGCACTTATGGAGGCACTTTCCAGCAGTTCAGAGCA 8739
QY 1293 GAACGATGATACATAGGCGCTTCAAGAGGAAATTTGAAGAACTTGAAGAACTTGAAGAACT 1352
Db 8740 GAACGATGATACATAGGCGCTTCAAGAGGAAATTTGAAGAACTTGAAGAACTTGAAGAACT 8799
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTTGGAGGACTAGAGAACT 1412
Db 8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTTGGAGGACTAGAGAACT 8859
QY 1413 CTACAGGAGCCAGAGAGCTGCTCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCT 1472
Db 8860 CTACAGGAGCCAGAGAGCTGCTCTGAGGAGAGAGCCAGAAATGTCACCTCGGCTTCT 8919
QY 1473 ACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAATTTGAACCTGCACTCCGCTGA 1532
Db 8920 ACGAAAGCAGGCTGAGGAGTCAATACTAGTGGGAAAATTTGAACCTGCACTCCGCTGA 8979
QY 1533 CTGGCAGAGAAAATAGATGAGACCTTTGAAGAACTTCAAGAACTTCAAGAGCCAGGGA 1592
Db 8980 CTGGCAGAGAAAATAGATGAGACCTTTGAAGAACTTCAAGAACTTCAAGAGCCAGGGA 9039
QY 1593 TGAGCTGGAACCTCAAGCTCGGCAAGCTGAGTGTATCAAGGATCCTGCGAGCCCGTGGG 1652

9040 TGAGCTGGACCTCAAGCTGGCCAGAGCTGAGTGTATCAAGGATCTGGCAGCCGCTGG 9099
1653 CGATCTCCCTCATGACTCTCTCCAGATCACTCGAGAAAGTCAAGGACCTTCGAGGAGA 1712
9100 CGATCTCCCTCATGACTCTCTCCAGATCACTCGAGAAAGTCAAGGACCTTCGAGGAGA 9159
1713 AATTGGCCCTCTGAAAGAGAAAGTGAAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCAC 1772
9160 AATTGGCCCTCTGAAAGAGAAAGTGAAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCAC 9219
1773 TTTGGCATTGAGCTCTCAGCTGATTAACCTCAGCACTCTGGAAGACCTGAACACCAAGATG 1832
9220 TTTGGCATTGAGCTCTCAGCTGATTAACCTCAGCACTCTGGAAGACCTGAACACCAAGATG 9279
1833 GAAGCTTTCTGAGCTGGCCGCTCGAGAGCCAGTCAAGGAGCTGCAATGAAGCCCAAGGGA 1892
9280 GAAGCTTTCTGAGCTGGCCGCTCGAGAGCCAGTCAAGGAGCTGCAATGAAGCCCAAGGGA 9339
1893 CTCTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTCAGGCTCTGCAAGGCTCTGGGAGAGC 1952
9340 CTCTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTCAGGCTCTGTCAGGCTCTGGGAGAGC 9399
1953 CATCTCGCAACAAAGTGCCTTACTATATCAACCAAGAGCTCAAA 2001
9400 CATCTCGCAACAAAGTGCCTTACTATATCAACCAAGAGCTCAAA 9448

RESULT 3

JS-08-836-022A-10/c
Sequence 10, Application US/08836022A
Patent No. 6001557
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
APPLICANT: Wilson, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Chen, Shu-Jen
APPLICANT: Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/836,022A
APPLICATION NUMBER: US/08/836,022A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVPN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA

US-08-836-022A-10
Query Match 50.2%; Score 1004; DB 3; Length 19307;
Best Local Similarity 90.3%; Pred. No. 6.1e-302;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
Qy 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCCCTGGACCTGGAAAAGTTTCTTCCCTG 872
Db 6434 GGAAGAAACTCATAGATTACTGCAAGAGTTCCCTCTGGACCTGGAAAAGTTTCTTCCCTG 6375
Qy 873 GCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGAGTGTCTACCCGTAAAGGAAAGGCT 932
Db 6374 GATTACGGAAGAGAAACAACTGCCAATGTCTTACAGAGTGTCTCCCGTAAGGAGAGCT 6315
Qy 933 CTTAGAAAGATCCAAAGGAGTAAAGAGCTGATGAACAAATGCAAGACCTTCCAGGTGA 992
Db 6314 CTTAGAAAGATCCAAAGGAGTAAAGAGCTGATGAACAAATGCAAGACCTTCCAGGTGA 6255
Qy 993 AATTGAAGCTCACACAGATGTTTATCACACCTGGATGAAACAGCAAAATTCCTGAG 1052
Db 6254 AATTGAAGCTCACACAGATGTTTATCACATCTTGTGAAATGGCCAAAATTCCTGAG 6195
Qy 1053 ATCCCTGGAAGGTTCCGATGATGCTGTTTCAAAAGACGTTTGGATTAACATGAACCTT 1112
Db 6194 ATCCCTGGAAGGTTCCGATGATGCTGTTTCAAAAGACGTTTGGATTAACATGAACCTT 6135
Qy 1113 CAAGTGAAGTGAATTCGGAAAAGTCTCTCAACATTTAGTCCCTCATTTGGAAGCCAGTTT 1172
Db 6134 CAAGTGAAGTGAATTCGGAAAAGTCTCTCAACATTTAGTCCCTCATTTGGAAGCAAGTTT 6075
Qy 1173 TGACCACTGGAAGGCTCTGCACCTTTCTTGCAGGAACCTTCTGCTGTGGCTACAGCTGAA 1232
Db 6074 TGACCACTGGAAGGCTTTGCACTTTCTTTCAGGAACCTTCTTGTGTGGCTACAGCTGAA 6015
Qy 1233 AGATGATGAATTAAGCCGGCAGCAGCACTTATTTGAGGGGACTTTCCAGCAGTTCCAGAGCA 1292
Db 6014 AGATGATGAATTAAGCCGGCAGCAGCACTTATTTGAGGGGACTTTCCAGCAGTTCCAGAGCA 5955
Qy 1293 GAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAACCTTAAAGAACCTTGAATCATGAG 1352
Db 5954 GAATGATATACATAGGGCTTCAAGAGGGAATTTGAAACCTTAAAGAACCTTGAATCATGAG 5995
Qy 1353 TACTCTTGAGACTGTACGAATATTTCTGAAGAGGACCTTTGGAAGGACTAGAGAACT 1412
Db 5894 TACTCTTGAGACTGTGAGAATATTTCTGAAGAGGACCTTTGGAAGGACTAGAGAACT 5835
Qy 1413 CTACCAAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGCCAGAAATGTCTACTCGGCTTCT 1472
Db 5834 CTACCAAGGAGCCAGAGAACTGCTCTCTGAGGAGAGCTCAGAAATGTCTACTCGGCTTCT 5775
Qy 1473 ACAGAAAGCAGGCTGAGGAGGTTCAATCTGAGTGGGAAAAATTGAACCTGCATCCGCTGA 1532
Db 5774 ACAGAAAGCAGGCTGAGAGGTTCAACGCTGAATGGGACAAATTTGAACCTGCCTGCTGAG 5715
Qy 1533 CTGCAAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGCCAGGA 1592
Db 5714 TTGCAAGAGAAAAATAGATGAGACTTTGAAAGACTCCAGGAACTTCAAGAGCTCCGGA 5555
Qy 1593 TGAGCTGGACCTCAAGCTGGCCAAAGCTGAGGAGTCAAGGGATCTCTGGCAGCCGCTGG 1652
Db 5654 TGAATGGACCTCAAGTTGGCCAAAGCTGAGGATCAAGGGATCTCTGGCAGCCGCTGG 5595
Qy 1653 CGATCTCTCTTACTCTCTCTCAGAGTCACTCTGAGAAAGTCAAGGACTTCCAGGAGA 1712
Db 5594 GGATCTCTCTTACTCTCTCTCAGAGTCACTCTGAGAAAGTCAAGGACTTCCAGGAGA 5535
Qy 1713 AATTGGCCTCTGAAAGAGAAAGTGAAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCAC 1772
Db 5534 AATTGGCCTCTTAAAGAGAAATGTCAATCGTGTCAATGAATTTGACATCAGCTGACCAC 5475
Qy 1773 TTTGGGATTCAGCTCTCAAGGATTAACCTCAGCACTCTCGAAGACCTGAAACCAAGATG 1832
Db 5474 ACTGGGCATTGAGCTCTCACTTATAACCTCAGCACTTTGGAAGATCTGAATACCAGATG 5415

1833 GAAGTCTTCAGGTGGCCGTCGAGGACCGAGTCAGGACGCTGATCAAGCCACACAGGGA 1892
5414 GAGGCTTCTACAGGTGGCTGTGGAGGACCGGTGTCAGACAGCTGATGAAGCCACACAGGGA 5355
1893 CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTCTGTCCAGGGTCCCTGGGAGAGAGC 1952
5354 CTTTGGTCTGCACTCCAGCACTTCTTCCACTTCAGTTTCAGGGTCCCTGGGAGAGAGC 5295
1953 CATCTGCGCAACAAAGGCGCTACTATATCAACACAGAGACTCAAAAC 2000
5294 CATCTCAGCAACAAAGGCGCTACTATATCAACACAGAGAGAGAGC 5247

RESULT 4
US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; Wilson, James M.
; Fisher, Krishna J.
; Chen, Shu-Jen
; Waitman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; Methods of Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,048A
; FILING DATE: 21-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,022
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNPVN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 50.2%; Score 1004; DB 3; Length 19307;
Best Local Similarity 90.3%; Pred. No. 6.1e-302;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

813 GGTACCTACTCATAGATTACTCAACAGTTCCCTGGACCTGGAAAGTTTCTTGCCTG 872
6434 GGAAGAACTCATAGATTACTGAGCAGTTCCCTGGACCTGGAGAGTTTCTTTCCTG 6375
873 GCTTACAGAGCTGAACAACTCCCACTGCTCTACAGGATGCTACCGTAAGGAAGGCT 932

6374 GATTACGGAAGCAGAAACAACTGCGCAATGTCTCTACAGACGCTTCCCGTAAGGAGAAGCT 6315
933 CTTAGAAAGCTCCCAAGGAGTAAAGAGCTGATGAACAAATGGCAAGACCTCCCAAGGTGA 992
6314 CTTAGAAAGCTCCCAAGGAGTCAAGAGCTGATGAACCAATGGCAAGATCTCCCAAGGAGA 6255
993 AATTGAAGCTCACACAGATGTTTATCAAACTGGATGAAAAACAGCCAAAAATCTCTGAG 1052
6254 AATTGAAGCTCACACAGATATCTATCACAATCTTGTGAAAAATGGCAAAAAATCTCTGAG 6195
1053 ATCCCTGGAAAGTTCGGATGATGAGTCTCTGTAACAAAGAGCTTTGGATATCAATGAACCTT 1112
6194 ATCCCTGGAAAGTTCGGATGAGACCCCTCTGTTACAAAGAGCTTTGGATATCAATGAATTT 6135
1113 CAAGTGGAGTGAACCTTCGAAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAAGCCAGTTC 1172
6134 CAAGTGGAGTGAACCTTCAGAAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAAGCAAGTTC 6075
1173 TGACCAAGTGGAAAGCTCTGCACTTCTCTGACAGAACTTCTGTTGTTGCTGCTACAGCTGAA 1232
6074 TGACCAAGTGGAAAGCTTTCGATCTTCTCTGAGAACTTCTGTTGTTGCTGCTACAGCTGAA 6015
1233 AGATGATGAATTAAGCCGCGCAGGCACTTATTGGAGGCGCACTTTCAGCAGTTCAGAAACA 1292
6014 AGATGATGAATTAAGCCGCGCAGGCACTTATTGGAGGCGCACTTTCAGCAGTTCAGAAACA 5955
1293 GAACGATGATACAGGAGCTTCAAGAGGGAATTAAGAACTTAAGAACTGTAATCATCAG 1352
5954 GAATGATATACATAGGAGCTTCAAGAGGGAATTAAGAACTTAAGAACTGTAATCATCAG 5895
1353 TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCTTTCAGAAAGTTCAGAAAGTTCAG 1412
5894 TACTCTTGAGACTGTGAGAAATTTCTGACAGAGCAGCTTTCAGAAAGTTCAGAAAGTTCAG 5835
1413 CTACCAAGGAGCCAGAGAGCTGCTCTCTGAGAGAGAGCCAGAAATGTCACCTCGCTTCT 1472
5834 CTACCAAGGAGCCAGAGAGCTGCTCTCTGAGAGAGAGCTCAGAAATGTCACCTCGCTTCT 5775
1473 AGAAAGAGAGCTGAGGAGCTCAATCTAGTGGGAAAAATTAAGAACTGTAATCATCAG 1532
5774 ACAGAGAGAGCTGAGGAGCTCAAGCTGATGAGGAGAAATTAAGAACTGTAATCATCAG 5715
1533 CTGGCAGAGAAAAATAGATGAGACCTTTGAAAGCTTTCAGAAAGTTCAGAAAGTTCAGAAAG 1592
5714 TTGGCAGAGAAAAATAGATGAGACCTTTGAAAGCTTTCAGAAAGTTCAGAAAGTTCAGAAAG 5655
1593 TGAGCTGGACCTCAAGCTGCGGCAAGCTGAGTGTCAAGGATCTCTGACAGCCCTGGG 1652
5654 TGAATGGACCTCAAGTTCGCGCAAGCTGAGTGTCAAGGATCTCTGACAGCCCTGGG 5595
1653 CGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTTCGAGGAGA 1712
5594 GGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTTCGAGGAGA 5535
1713 AATTGGCCTCTGAAAGAGAGCTGAGCCACCTGATGAGCTTGTCTGCGCAGCTTACCAAC 1772
5534 AATTGCACTCTTTAAAGAGAGTGTCAATCGTCAATGACCTTGCACATCAGCTGAGCCAC 5475
1773 TTTGGCATTTCAGCTCTCACCGTATTAACCTCAGCACTCTGGAAGACCTGAAACCCAGATG 1832
5474 ACTGGGCATTTCAGCTCTCACCTTATTAACCTCAGCACTTTCGGAAGATCTGATACCAAGATG 5415
1833 GAAGCTTTCGAGGTGGCGTTCGAGAGCCAGTTCAGGAGCTGCGATGAGCCCAAGGAGA 1892
5414 GAGGCTTCTACAGGTGGCTGTGGAGAGCCGTGTGACAGAGCTGCTGAGAGCCCAAGGAGA 5355
1893 CTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGAGC 1952
5354 CTTTGGTCTGCACTCCAGCACTTCTTCCACTTCAGTTTCAGGGTCCCTGGGAGAGAGC 5295
1953 CATCTGCCCAACAAAGTGCCTTACTATATCAACACAGAGACTCAAAAC 2000
5294 CATCTCAGCAACAAAGTGCCTTACTATATCAACACAGAGAGAGAGC 5247

RESULT 5

US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(6037)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7
Query March 19 88; Score 397; DB 4; Length 6045;
Best Local Similarity 59.3%; Pred. No. 6.8e-113; Indels 9; Gaps 1;
Matches 698; Conservative 0; Mismatches 470;
833 TGCACAGTTCCTCCCTGAGACTGCGGAAAGTTCTTGTGCTGCTTACAGAGCTGGAACAA 892
3069 TGCAGGCTCTCGCAGAGATCTGGAAGTCTCTGAGTGATGATGATGATGATGATGATGAT 3128
893 CTGCCAATGCTTACAGAGTCTACCGTGAAGGAAGGCTCTAGAGACTCCAAAGGAG 952
3129 CAGTGAATGTGCTTGTGATGCTCTCATCGGAGAGATGCTCTTCCAGGATAGTATCTTGG 3188
953 TAAAGAGCTGATGAACAATGGCAAGACTCCAAAGTGAAATTTGAAGCTTCACACAGATG 1012
3189 CCAGGAACTCAACACAGAGATGAGGACATCCAGGAGAAATTTGATGCCCAATGACA 3248
1013 TTTATCAAACTCGATGAAACAGCAAAATCTGAGATCCCTGGAAGGTTCCGATG 1072
3249 TATTTAAAGCATTCAGGAAACAGGAGAGATGGTAAAGGTTTGGGAAATCTTGAAG 3308
1073 ATGAGTCTCTTCAAGAGCTTTGGTAACTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1132
3309 AGGCTATGCTTCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3368
1133 AAAAGTCTCTCAACATTAGTTCCTTGGAGCCAGTCTTGACCAAGTGAAGCGTCTGC 1192
3369 CAAATCTGCTAGCATCAGGCCCCATTTGGAGGCCAGCGCTGAGAGTGAACAGGTTGC 3428
1193 ACCTTCTCTGAGGAATCTTGTGTGCTTACAGCTGAAAGATGATGATTAAGCCGGC 1252
3429 TGATGCTCTTGAAGAACTGATCAATGGCTGAATATGAAGATGAAGAGCTTAAGAAAC 3488
1253 AGGCACCTTATGGAGGCGACTTTCAGCAGTTCCAGAGCTTCAAGAGCAAGATGATAGGCGCT 1312
3489 AATGCCCTTATGGAGGAGATGTTCCAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAG 3548

RESULT 6

US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
1313 TCAAGAGGGAATTGAAAACCTAAAGAACCTGTAAATCATAGTACTCTTTGAGACTGTACGAA 1372
3549 TGAGACGGAGTTAAAGGAGAAAGATATTCTGTCTCTGAATGTGTGACGAGCCCGAG 3608
1373 TATTTCTGACAGAGCAGCCCTTT-----GGAAGGACTAGAGAACTCTACCAAGAGC 1423
3609 TTTTCTTGGCTGATCAGCAATTTGAGGCCCTTGAAGAGCCCAAGAGAACTTACATCAA 3668
1424 CCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGCTTTCTAGAAAGCAGG 1483
3669 AAAAGATTAACCTCTGAGGAGAGAGCCCAAGAAATGTCGCAAGCCATGCGCAACAGT 3728
1484 CTGAGGAGCTCAATCTGAGTGGGAAAAATTAACCTGCACTCGCTGACTGGGAGAGAA 1543
3729 CTTCTGAAGTCAAAAGAAAAATGCGGAAAGTCTAAATGCTGTAACTAGCAATTTGGCAAGC 3788
1544 AATAGATGAGACCCCTTGAAGAGCTCCAGGAACCTTCAAGAGGCCCAAGAGTGGTGGACC 1603
3789 AAGTGGACAGGCAATTGGAGAACTTCAGAGACCTGCAAGGAGCTTATGATGACCTGGAGC 3848
1604 TCAAGCTGCGCAAGCTGAGTGATCAAGGATCTCTGGAGCCCGTGGGCGATCTCTCA 1663
3849 CTGACATGAGAGGAGGAGAGTCCGTGCGGAATGCTGGAAGCCCGTGGGAGACTTACTCA 3908
1664 TTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTCGGCTC 1723
3909 TTGACTCTCTCGAGATCACTTGAAGAAATCATGGGCAATTAGAGAGAAATTCACCAA 3968
1724 TGAAGAGAACGTGAGCCACCTCAATGACCTTGTGCGGAGCTTACCCTTTGGGCAATTC 1783
3969 TCAACTTTAAAGTTAAAGACGGTGAATGATTTATCCAGTCACTGTCTCCACTTTGACCTGC 4028
1784 AGCTCTACCCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGAGTGAAGCTTCTGC 1843
4029 ATCCCTCTCTAAAGATGTCTGCCAGCTAGATGACCTTAATATCGATGAAACTTTTAC 4088
1844 AGTGGCCGCTGAGGAGCCGAGTCAAGGAGCTGCAATGAAGCCCAAGAGCTTTGGTCCAG 1903
4089 AGTTCCTGTGATGATGCTGCTTAAACAGCTTCAGGAAGCCCAAGAGATTTTGGACAT 4148
1904 CATCTCAGCACTTTCTTCCAGCTCTGTCAGGCTCTGTCAGGCTCCCTGGAGAGAGCCATCTCGCAA 1963
4149 CCTCTCAGCACTTTCTCTCTACGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4208
1964 ACAAGTCCCTACTATATCAACCAAGAGACTCAAAAC 2000
4209 ATAAAGTCCCTATTATCAATCAACATCAACACAGAC 4245

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
; US-09-091-501B-9

Query Match          19.8%; Score 397; DB 4; Length 10320;
Best Local Similarity 59.3%; Pred. No. 9.9e-113;
Matches 698; Conservative 0; Mismatches 470; Indels 9; Gaps 1;

QY 833 TGAACAGTTCGCCCTGGACCTGGAAAAGTTCTTGGCTTGGCTTACAGAAGCTGAAACAA 892
DB 833 TGAACAGTTCGCCCTGGACCTGGAAAAGTTCTTGGCTTGGCTTACAGAAGCTGAAACAA 892
QY 834 CTGCGCAATGCTCTACAGGATGCTACCGGTAAGGAAAGGCTCTTGAAGACTCCAGGGAG 952
DB 834 CTGCGCAATGCTCTACAGGATGCTACCGGTAAGGAAAGGCTCTTGAAGACTCCAGGGAG 952
QY 835 CAGTGAATGCTTGTGGATGCTCTCATCGGGAATGCTCTTCCAGATAGTATCTTGG 7463
DB 835 CAGTGAATGCTTGTGGATGCTCTCATCGGGAATGCTCTTCCAGATAGTATCTTGG 7463
QY 953 TAAAGAGCTGATGAACAATGCGAAGACTCCAAAGGTGAATTTGAAGCTCACACAGATG 1012
DB 953 TAAAGAGCTGATGAACAATGCGAAGACTCCAAAGGTGAATTTGAAGCTCACACAGATG 1012
QY 7464 CCAGGGAATCTCAACAGCAGATGCGAGGATCCAGGACAGAAATGATGCCACATGACA 7523
DB 7464 CCAGGGAATCTCAACAGCAGATGCGAGGATCCAGGACAGAAATGATGCCACATGACA 7523
QY 1013 TTATACAACTGGATGAAGACAGCAAAATCTTGATCCCTGGAGTTCCTGGATG 1072
DB 1013 TTATACAACTGGATGAAGACAGCAAAATCTTGATCCCTGGAGTTCCTGGATG 1072
QY 7524 TATTTAAAGCATTTGACGGAACAGCAAGATGTTAAAGCTTTGGGAATTTCTGAAG 7583
DB 7524 TATTTAAAGCATTTGACGGAACAGCAAGATGTTAAAGCTTTGGGAATTTCTGAAG 7583
QY 1073 ATCAGTCTCTTACAAAGACGTTTCGGATAAATGAATCTCAAGTGGAGTGAATTCGGA 1132
DB 1073 ATCAGTCTCTTACAAAGACGTTTCGGATAAATGAATCTCAAGTGGAGTGAATTCGGA 1132
QY 7584 AGCTACTATGCTTCAACATCGACTGATGATGATGACCAAGATGGAATGACTTAAAG 7643
DB 7584 AGCTACTATGCTTCAACATCGACTGATGATGATGACCAAGATGGAATGACTTAAAG 7643
QY 1133 AAAAGTCTCTCAACATGATGCTCCATTTGGAAGCAGTGTCTGACAGTGGAGGCTTGC 1192
DB 1133 AAAAGTCTCTCAACATGATGCTCCATTTGGAAGCAGTGTCTGACAGTGGAGGCTTGC 1192
QY 7644 CAAAATCTGTAGCATCAGGGCCATTTGGAGCCAGGCTGAGAAGTGGAAACAGGTTTC 7703
DB 7644 CAAAATCTGTAGCATCAGGGCCATTTGGAGCCAGGCTGAGAAGTGGAAACAGGTTTC 7703
QY 1193 ACCTTCTCTGAGGAACTTCTGCTGGCTACAGCTGAAAGATGATGAATTAAGCCGCG 1252
DB 1193 ACCTTCTCTGAGGAACTTCTGCTGGCTACAGCTGAAAGATGATGAATTAAGCCGCG 1252
QY 7704 TGATGCTTTAGAGAACTGATCAATGCTGATGATGATGATGATGATGATGATGATG 7763
DB 7704 TGATGCTTTAGAGAACTGATCAATGCTGATGATGATGATGATGATGATGATGATG 7763
QY 1253 AGGACCTATTTGAGGCGACTTTCCAGAGTTCAGAGCAGAACAGATGATGATGATGATG 1312
DB 1253 AGGACCTATTTGAGGCGACTTTCCAGAGTTCAGAGCAGAACAGATGATGATGATGATG 1312
QY 7764 AAATGCTATTTGAGGAGATGTTCCAGGCTTACAGCTCCAGTATGACCAATTTGAAGGCC 7823
DB 7764 AAATGCTATTTGAGGAGATGTTCCAGGCTTACAGCTCCAGTATGACCAATTTGAAGGCC 7823
QY 1313 TCAAGAGGGAATTTGAACCTTAAGACCTGTAATCATGATGATGATGATGATGATGATGATG 1372
DB 1313 TCAAGAGGGAATTTGAACCTTAAGACCTGTAATCATGATGATGATGATGATGATGATGATG 1372
QY 7824 TGACAGGGGATTTAAAGGAGAAAGATATTTCTGCTGAAATGATGATGATGATGATGATGATG 7883
DB 7824 TGACAGGGGATTTAAAGGAGAAAGATATTTCTGCTGAAATGATGATGATGATGATGATGATG 7883
QY 1373 TATTTCTGACAGAGCAGCTTTT-----GGAAGGACTAGAGAACTCTTACAGGAGC 1423
DB 1373 TATTTCTGACAGAGCAGCTTTT-----GGAAGGACTAGAGAACTCTTACAGGAGC 1423
QY 7884 TTTTCTTGGCTGATCAGGCAATTTGAGGCCCTTGAAGCAGCAAGAACCTTACATCAA 7943
DB 7884 TTTTCTTGGCTGATCAGGCAATTTGAGGCCCTTGAAGCAGCAAGAACCTTACATCAA 7943
QY 1424 CCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGATGTCTACTGGCTTCTACGAAGCAGG 1483
DB 1424 CCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGATGTCTACTGGCTTCTACGAAGCAGG 1483
QY 7944 AAACAGAAATTAATCTCTGAGGAGAGAGCCAAAGATTTGCCAAAGCATGCGCAACAGT 8003
DB 7944 AAACAGAAATTAATCTCTGAGGAGAGAGCCAAAGATTTGCCAAAGCATGCGCAACAGT 8003
QY 1484 CTGAGGAGGTCAATTAATCTGAGTGGGAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAA 1543
DB 1484 CTGAGGAGGTCAATTAATCTGAGTGGGAAAATTTGAACCTGCACTCCGCTGACTGGCAGAGAA 1543
QY 8004 CTCTGAGTCAAGAAAATTTGGAAGTCTAAATGCTGTAATGCTGTAATGCTGTAATGCTGTAATG 8063
DB 8004 CTCTGAGTCAAGAAAATTTGGAAGTCTAAATGCTGTAATGCTGTAATGCTGTAATGCTGTAATG 8063
QY 1544 AAATAGATGAGACCTTTGAAAGACTCCAGGAATTTCAAGAGGCCAGGATGATGATGATGATG 1603
DB 1544 AAATAGATGAGACCTTTGAAAGACTCCAGGAATTTCAAGAGGCCAGGATGATGATGATGATG 1603
QY 8064 AAGTGGACAAGGCAATTTGGAAGAACTCAGAGACCTGAGGAGGATGATGATGATGATGATGATG 8123
DB 8064 AAGTGGACAAGGCAATTTGGAAGAACTCAGAGACCTGAGGAGGATGATGATGATGATGATGATG 8123
QY 1604 TCAAGCTGCCCGAGCTGAGTGTATCAAGGATCTTGGCAGCCGCTGGCGGATCTCTCA 1663
DB 1604 TCAAGCTGCCCGAGCTGAGTGTATCAAGGATCTTGGCAGCCGCTGGCGGATCTCTCA 1663
QY 8124 CTGACATGAAGGAGGAGATGCTCGTGGGAATGGCTTGGGAAGCCCTGGGAGACTTACTCA 8183
DB 8124 CTGACATGAAGGAGGAGATGCTCGTGGGAATGGCTTGGGAAGCCCTGGGAGACTTACTCA 8183
```

RESULT 7

```
US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Rattus sp.
; US-09-091-501B-5
```

```
Query Match          4.0%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. 7.9e-15;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

```
QY 465 CCTAAAACGCCAAGTACAAACAATAAGGTGCTTCAAGAAGATCTAGAACAAACAAGT 524
DB 16 CCTGCAAAACCTGCTTGAAGAAATATAAAGTTTGAAGTGAAGTCTGAGCTGAGCAGGT 75
QY 525 CAGGCTCAATTTCTCACTCAATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 584
DB 525 CAGGCTCAATTTCTCACTCAATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 584
QY 76 GAAGGTGAATTCCTTAATCATATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 135
DB 76 GAAGGTGAATTCCTTAATCATATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 135
QY 585 AACTGCTGCTTTGGAGAAACAATTAAGGTATTGGGAGATCGATGGGCAACAATCTGTAG 644
DB 136 CACAGCTGTTTGGAGATCAGTTTACAGAACTGGGTGAGCGCTGGACACTGTATGCGG 195
QY 645 ATGGA 649
DB 196 CTGGA 200
```



```
RESULT 8
US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091.501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-091-501B-4

Query Match      3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.4e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 453 TGATCTTTGAGACCTAAACGCCCAAGTACACATAGGTGCTTCAAGAGATCTAGA 512
Db 4 TGACCTGCCCTCCTCGAGAGTGTCTTCAAGAACATAAAAGTTTCAAAATGACCTTGA 63
QY 513 ACAAGAACAAAGTCAGGTCAATTTCTCACTACATGTTGGTGTAGTGAATCTAG 572
Db 64 AGCTGAACAGTGAAGTAAATTCCTTAACATCATGTTGGTGTAGTGAATCTAG 123
QY 573 TGGAGATCAGCACTGCTCTTGGAGAACAACTTAAGTATTTGGAGATCGATGGGC 632
Db 124 TGGGGAGTGCACACGCTTCTTGAAGATCAGTTACAGAACTGGTGGCGCTGGAC 183
QY 633 AAACATCTGTAGATGGA 649
Db 184 AGCTGTATGCCGTGGA 200

RESULT 9
US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091.501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 6
; LENGTH: 200
; TYPE: DNA
US-09-091-501B-6

Query Match      3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.4e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 453 TGATCTTTGAGACCTAAACGCCCAAGTACACATAGGTGCTTCAAGAGATCTAGA 512
Db 4 TGACCTGCCCTCCTCGAGAGTGTCTTCAAGAACATAAAAGTTTCAAAATGACCTTGA 63
QY 513 ACAAGAACAAAGTCAGGTCAATTTCTCACTACATGTTGGTGTAGTGAATCTAG 572
Db 64 AGCTGAACAGTGAAGTAAATTCCTTAACATCATGTTGGTGTAGTGAATCTAG 123
QY 573 TGGAGATCAGCACTGCTCTTGGAGAACAACTTAAGTATTTGGAGATCGATGGGC 632
Db 124 TGGGGAGTGCACACGCTTCTTGAAGATCAGTTACAGAACTGGTGGCGCTGGAC 183
QY 633 AAACATCTGTAGATGGA 649
Db 184 AGCTGTATGCCGTGGA 200
```

```
; ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match      3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.4e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 453 TGATCTTTGAGACCTAAACGCCCAAGTACACATAGGTGCTTCAAGAGATCTAGA 512
Db 4 TGATCTGAAATCTCTACAAAAGCTGCTAGAAAGACATAAAAGTTTGCAAAAGTGAATCTTGA 63
QY 513 ACAAGAACAAAGTCAGGTCAATTTCTCACTACATGTTGGTGTAGTGAATCTAG 572
Db 64 GGCTGAACAGGTGAAAGTAAATTCACAACTCAGTGTGCTCAATTTGTTGATGAAAACAG 123
QY 573 TGGAGATCAGCACTGCTCTTGGAGAACAACTTAAGTATTTGGAGATCGATGGGC 632
Db 124 TGGTGAGAGCGCTACAGCTATCTCTAGAAAGACCACTTACAGAACTTTGGTGAGCGCTGGAC 183
QY 633 AAACATCTGTAGATGGA 649
Db 184 AGCAGTATGCCGTGGA 200

RESULT 10
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; CLONE: PTZgpt-F1s
; IMMEDIATE SOURCE:
US-08-232-463-14
```



```
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (300001)..(315000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (315001)..(330000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
```

```
NAME/KEY: misc feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
```

Query Match 2.2%; Score 44; DB 4; Length 1230025;

Best Local Similarity 53.5%; Pred. No. 0.39;

Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
QY 242 AATTATCAGAGATGAAGAAGTCAAGTACAGAGCAGATGTAATCTCTAAATTCAGAT 301
Db 656477 AATTAGAAGAAGAAGAAGAAGAAATTGAGGATATCAAGACTCAGATACAAAT 656418
QY 302 GGGARTGCTCAGGGTAGCTAGCATGGGAAACAAAGCAATTTATAGATTTTATAGG 361
Db 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCCTAGGCAAGCAATTTATG 656358
QY 362 ATCTCCAGATCAGAAACTGAAGAGTTGAATGCTACTGGCTAACAAAACAGA 413
Db 656357 CAPTTAAGCAGAAAAAACTAAAGCTTCTTAAGAGACGCGCTGGGAATAGA 656306
```

RESULT 13

US-09-107-532A-1186

; Sequence 1186, Application US/09107532A

Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Atinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1186:
SEQUENCE CHARACTERISTICS:
LENGTH: 1179 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1179
SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186
Query Match 2.2%; Score 43.4; DB 4; Length 1179;
Best Local Similarity 50.7%; Pred. No. 0.0047;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 371 ATCAGAACTGAAGAGTTCAATGACTGCTAACAAAAACAGAGAAAGAAACAGGAAAA 430
DB 602 AGCAAGATAAAGATTGATCGATCAGACAAAGAAATAGATACGATCGAGGAA 661
QY 431 TGAGAGAGAGCCTCTTGACCTGATCTTGAAGACCTTAAACGCCAGTACAAACATA 490
DB 662 TTGTAGAAGTGCTCGTTGGAGGCGTTCCAGCTGGATTAGGAAGCTACGTTACATGGGACA 721
QY 491 AGGTGCTTCAAGAAGATCTAGAACAGAAACAGTACAGGCTCAATTCTTCACTCACATGG 550
DB 722 CGAAGCTAGTAGCAAAATCGCAACAGCTGTGGTTAGTATCAATGCCITTAAGCGGTAG 781
QY 551 TGGTGGTAGTGAATGATCTAGTGG 575
DB 782 AATTGGGGTCGGATTCACTTCTGG 806
RESULT 14
US-09-620-312D-69

Sequence 69, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aigong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunding
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP28
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 69
LENGTH: 1690
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (128) .. (1522)
US-09-620-312D-69
Query Match 2.1%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.0092;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
QY 1569 CCAGGAACCTCAGAGGCCACCGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGTGAT 1628
DB 55 CAAGGAGTTGCACCGTGGCGCACCGCTGAGCGAGCTGGCATGGTTTCAGGAGCG 114
QY 1629 CAAGGGA---TCCTGGCAGCCCGTGGCGCATCTCTTCATTGACTCTTCCAAAGATCACCT 1685
DB 115 GCTGCCACTGGCCATGCAGACAGAGCGAGCAACGGTTTGCAGCGCGTCCAGCAGCACAT 174
QY 1686 CGAGAAAGTCAAGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAACTGAGCCACCT 1745
DB 175 CAAAAAGAACAGGGGCTCGCGCGGAGATTCGAGCGCATGGCCGCTGGAGAGGT 234
QY 1746 CAATGACCTTGTCTGCGCAGCTTACCACTTTGGSCATTCAGCTCTCACCGTAACTCAG 1805
DB 235 GCTGGAGCGCGCGCGCGCTGGCGTCTGGCGAGCCCGGAGGACAGAGGAGTGGCGCG 294
QY 1806 CACTCTGGAAGAGCTGAACACAGATGGAAGCTTCTGAGGTTGGCGTGGAGACCGAGT 1865
DB 295 GGGCTTGGACAGCTGCAGAGCGCTGGCGCGGAGTGGCGCGGAGTGGCGCGGAGTGGCA 354
QY 1866 CAGCAGCTGCATGAAGCCCAACAGGG 1891
DB 355 GCAGGTCTGACGCGCGCTTCCAGG 380
RESULT 15
US-09-368-590-1
Sequence 1, Application US/09368590
Patent No. 6187563

GENERAL INFORMATION:
APPLICANT: Solimena, Michela
TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
FILE REFERENCE: 101918-200 (OCR-941)
CURRENT APPLICATION NUMBER: US/09/368,590
CURRENT FILING DATE: 1999-08-04
EARLIER APPLICATION NUMBER: 60/095,657
EARLIER FILING DATE: 1998-08-07
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 7812
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(6879)
NAME/KEY: unsure
LOCATION: (100)...(102)
NAME/KEY: unsure
LOCATION: (1021)...(1023)
NAME/KEY: unsure
LOCATION: (2266)...(2268)
JS-09-368-590-1

Query Match 2.1%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.027;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

2y	1569	CGAGGACTTCAGAGGCCACGGATGAGTGGACCTCAGCTGCCCGCAGCTGAGGTGAT	1628
2b	3639	CAAGGAGTTGCACCGAGTGGCGCACACCTGGACGAGCTGGGTTCAGGACG	3698
2y	1629	CAAGGGA--TCTGGCAGCCGTGGCGCATCTCCTCATTTGACTCTCTCCAGATCACCT	1685
2b	3699	GCTGCCACTGGCATGCACAGAGCGAGGCAACGTTTTCAGGCGGTCCAGCACACAT	3758
2y	1686	CGAGAAAGTCAGGACACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAGCTGAGCCAGT	1745
2b	3759	CAAAAGAACAGAGGCTCGCGCGGAGATCCAGGCGCATGGGCCCGCCCTGGAGAGGT	3818
2y	1746	CAATGACCTTGTTCGCGAGCTTACCACTTTGGGCATTGAGCTCTCACCGTATAACCTCAG	1805
2b	3819	GCTGGAGCGCGCGGCGCGCTGGCTGCTGCGACCCGAGGAGGAGGAGTGGCGCG	3878
2y	1806	CACCTCTGGAGACCTGAACACCAAGATGGAAGCTTTGCAAGTGGCGCTGAGGACCGAGT	1865
2b	3879	GGCCCTGGAGACGCTGACAGAGCCCTGGGCGGAGCTGCGGAGGCTGCCGACGACGCA	3938
2y	1866	CAGGACGCTGCATGAAGCCACAGGG	1891
2b	3939	GCAGGTGCTGGACGCGCCTTCAGG	3964

Search completed: February 2, 2004, 06:15:49
Job time : 127.697 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 475.292 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000

Perfect score: 2001
Sequence: 1 ggcagttcattgatggagag.....tcaaccacgagactcaaac 2001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2001	100.0	3999	24	Human dystrophin m
2	2001	100.0	4966	24	Adeno-associated v
3	2001	100.0	4990	24	Adeno-associated v
4	1709	85.4	3858	24	Human dystrophin m
5	1709	85.4	4825	24	Adeno-associated v
6	1709	85.4	4848	24	Adeno-associated v
7	1709	85.4	5060	24	Adeno-associated v
8	1662.8	83.1	4182	24	Human dystrophin m

9	1662.8	83.1	5149	24	AAD37255	Adeno-associated v
10	1283	64.1	5462	24	ABK81999	DNA encoding mini-
11	1182.6	59.1	5952	22	AAD06794	Human dystrophin g
12	1182.6	59.1	8689	24	ABK82000	DNA encoding mini-
13	1182.6	59.1	11058	24	AAD37229	Human dystrophin p
14	1182.6	59.1	11241	24	ABK82005	CDNA encoding huma
15	1182.6	59.1	11443	24	AAD82002	DNA encoding mini-
16	1182.6	59.1	12923	10	ABN90338	Sequence of human
17	1182.6	59.1	13957	24	ABT10904	Human breast cance
18	1182.6	59.1	13957	24	ABG69900	Human dystrophin g
19	1182.6	59.1	13957	24	ABN95786	Gene #2284 used to
20	1182.6	59.1	13957	24	ABK81959	CDNA encoding huma
21	1182.6	59.1	13977	24	ABK81959	Human bone remodel
22	1182	59.1	2169	24	AAD37232	Human dystrophin r
23	1182	59.1	3531	24	AAD37238	Human dystrophin m
24	1182	59.1	4498	24	AAD37258	Adeno-associated v
25	1180	59.0	5339	24	ABK81998	DNA encoding mini-
26	1013	50.6	3510	24	AAD37240	Human dystrophin m
27	1013	50.6	4476	24	AAD37259	Adeno-associated v
28	1004	50.2	13815	19	AAV18885	Mus musculus dyst
29	1004	50.2	13815	24	ABK81960	CDNA encoding mous
30	1004	50.2	13815	24	ABK81960	Mouse ischaemic co
31	1004	50.2	19307	17	AAI27558	Shuttle vector pad
32	835	41.7	1821	24	AAD37241	Human dystrophin r
33	727	36.3	5417	24	ABK81997	DNA encoding mini-
34	678.4	33.9	4402	21	AAZ48568	A rod shortened dy
35	678.4	33.9	4414	24	AAD37260	Adeno-associated v
36	678	33.9	1991	24	AAD37231	Human dystrophin N
37	677	33.8	1667	24	AAD37235	Human dystrophin N
38	666.4	33.3	3446	24	AAD37242	Human dystrophin m
39	618.8	30.9	4075	21	AAZ48569	A rod shortened dy
40	614.2	30.7	4402	21	AAZ48567	A rod shortened dy
41	577.8	26.9	3275	10	AAK97129	Partial sequence o
42	448	22.4	1434	24	AAD37243	Human dystrophin r
43	408.2	20.4	11096	24	ABK81962	CDNA encoding mous
44	397	19.8	6045	18	AAI74655	Utrrophin minigene.
45	397	19.8	6059	22	AAI84673	Nucleotide sequenc

ALIGNMENTS

RESULT 1
AAD37234
ID AAD37234 standard; DNA; 3999 BP.

XX AAD37234;

AC AAD37234;

DT 21-AUG-2002 (first entry)

DE Human dystrophin minigene delta3990.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

OS Homo sapiens.

PN WO200183695-A2.

XX 08-NOV-2001.

PF 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

PI Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular

PT

dystrophy comprises an N-terminal domain or modified N-terminal domain, PT
rod repeats, H1 and H4 domains and a cysteine rich domain of a
dystrophin gene -

Example 1; Page 46-47; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified
N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is human
dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge
CC H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 other;

Query Match		100.0%; Score 2001; DB 24; Length 3999;
Best Local Similarity		100.0%; Pred. No. 0;
Matches 2001; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
QY	1	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAA 60
DB	1000	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAA 1059
QY	61	GTATTATCGTGCTCTTTCTGCTGAGGACACATTTGCAAGCACAAGAGAGATTTCTTAAT 120
DB	1060	GTATTATCGTGCTCTTTCTGCTGAGGACACATTTGCAAGCACAAGAGAGATTTCTTAAT 1119
QY	121	GATGTGGAAGTGGTGAAGACCAAGTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
DB	1120	GATGTGGAAGTGGTGAAGACCAAGTTCATCTCATGAGGGGTACATGATGGATTTGACA 1179
QY	181	GCCATFCAGGCGCGGTGGTGAATTTCTACAAATTTGGGAAGTACGATGTTGGAAACAGGA 240
DB	1180	GCCATFCAGGCGCGGTGGTGAATTTCTACAAATTTGGGAAGTACGATGTTGGAAACAGGA 1239
QY	241	AAATTATCAGAGATCAGAACTGAAGTGAAGTACAGAGCAGATCAATCTCTAAATTCAGA 300
DB	1240	AAATTATCAGAGATCAGAACTGAAGTGAAGTACAGAGCAGATCAATCTCTAAATTCAGA 1299
QY	301	TGGGAATGCTTCAGGGTAGCTAGCTAGCAAAACCAAGCAATTTACATAGAGTTTAATG 360
DB	1300	TGGGAATGCTTCAGGGTAGCTAGCAAAACCAAGCAATTTACATAGAGTTTAATG 1359
QY	361	GATCTCAGATCAGAACTGAAGAGTTCATGATGCTGCTACCAAAACAGAGAAAGA 420
DB	1360	GATCTCAGATCAGAACTGAAGAGTTCATGATGCTGCTACCAAAACAGAGAAAGA 1419
QY	421	ACAAGGAAATTCAGGAAGAGCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 480
DB	1420	ACAAGGAAATTCAGGAAGAGCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 1479
QY	481	CAACAAATAGGTGCTTCAAGAGATCTAGAACAGAACCAAGTCAAGGTCAATTTCTCTC 540
DB	1480	CAACAAATAGGTGCTTCAAGAGATCTAGAACAGAACCAAGTCAAGGTCAATTTCTCTC 1539
QY	541	ACTCACATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
DB	1540	ACTCACATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1599
QY	601	GAACTTTAAGGTATTTGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCG 660
DB	1600	GAACTTTAAGGTATTTGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCG 1659
QY	661	TGGGTTCTTTTACAGACCGCTGACCTAGCTCTCGACTGACCTATTTGGAGCTCT 720
DB	1660	TGGGTTCTTTTACAGACCGCTGACCTAGCTCTCGACTGACCTATTTGGAGCTCT 1719

QY	721	CCTACTCAGACTGTTTACTCTGTCACAACTGTGTTTACTTAAGGAACTGCCATCTCC 780
DB	1720	CCTACTCAGACTGTTTACTCTGTCACAACTGTGTTTACTTAAGGAACTGCCATCTCC 1779
QY	781	AACTAGAAATGCGCATCTCTTGTGATGTTGAGGTACTTACTAGTATTACTGCAACAG 840
DB	1780	AACTAGAAATGCGCATCTCTTGTGATGTTGAGGTACTTACTAGTATTACTGCAACAG 1839
QY	841	TTCCCTCTGGACCTGGAAGAGTTTCTTGCCTGGCTTACAGAGCTGGAACAACTGCCAAT 900
DB	1840	TTCCCTCTGGACCTGGAAGAGTTTCTTGCCTGGCTTACAGAGCTGGAACAACTGCCAAT 1899
QY	901	GTCTTACAGGATGCTACCCGTAAGAAAGGCTCTTAGAAGACTCCAGGGAGTAAAGAG 960
DB	1900	GTCTTACAGGATGCTACCCGTAAGAAAGGCTCTTAGAAGACTCCAGGGAGTAAAGAG 1959
QY	961	CTGATGAACATGCAAGACCTCAAGCTGAAATTTGAAGCTCACACAGATGTTTATCAC 1020
DB	1960	CTGATGAACATGCAAGACCTCAAGCTGAAATTTGAAGCTCACACAGATGTTTATCAC 2019
QY	1021	AACTGGATGAAACAGCCAAATAATCTCGATCCCTGGAGGTTCCGATGATGCGAGTC 1080
DB	2020	AACTGGATGAAACAGCCAAATAATCTCGATCCCTGGAGGTTCCGATGATGCGAGTC 2079
QY	1081	CTGTTACAAAGAGCTTTGGATTAACATGAACTTCAAGTGGAGTGAATTCGGAATAAGTCT 1140
DB	2080	CTGTTACAAAGAGCTTTGGATTAACATGAACTTCAAGTGGAGTGAATTCGGAATAAGTCT 2139
QY	1141	CTCAACATTTAGTCTCCATTTTGAAGCCAGTCTTGACCAAGTGAAGCGCTGCACTTTCT 1200
DB	2140	CTCAACATTTAGTCTCCATTTTGAAGCCAGTCTTGACCAAGTGAAGCGCTGCACTTTCT 2199
QY	1201	CTGCAAGAACTTCTGTTGGTCTACAGCTGAAGATGATTAAGTCGCGGAGACACCT 1260
DB	2200	CTGCAAGAACTTCTGTTGGTCTACAGCTGAAGATGATTAAGTCGCGGAGACACCT 2259
QY	1261	ATTGGAGGCGACTTTCCAGCAGTTCAGAAAGCAAGATGTACATAGGCGCTTCAAGAG 1320
DB	2260	ATTGGAGGCGACTTTCCAGCAGTTCAGAAAGCAAGATGTACATAGGCGCTTCAAGAG 2319
QY	1321	GAATGAAATCTAAAGAACTGTATCATGATGATCTTCTTGAGACTGTACGAATTTCTG 1380
DB	2320	GAATGAAATCTAAAGAACTGTATCATGATGATCTTCTTGAGACTGTACGAATTTCTG 2379
QY	1381	ACAGAGCAGCTTTTGAAGGACTTAGAAGAACTCTACAGAGGCCAGAGAGCTGCTCT 1440
DB	2380	ACAGAGCAGCTTTTGAAGGACTTAGAAGAACTCTACAGAGGCCAGAGAGCTGCTCT 2439
QY	1441	GAGGAGAGAGCCAGAAATGTCCTCGGCTTCTACGAAAGAGCGCTGAGGAGTCAATCT 1500
DB	2440	GAGGAGAGAGCCAGAAATGTCCTCGGCTTCTACGAAAGAGCGCTGAGGAGTCAATCT 2499
QY	1501	GAGTGGGAAATTTGAACTGCACTCGCTGCTGCGGAGGAAATAGATGAGAGCCCTT 1560
DB	2500	GAGTGGGAAATTTGAACTGCACTCGCTGCTGCGGAGGAAATAGATGAGAGCCCTT 2559
QY	1561	GAAAGACTCCAGGAATCTCAAGAGGCCACGGATGAGCTCAAGCTGCGGCCAAGCT 1620
DB	2560	GAAAGACTCCAGGAATCTCAAGAGGCCACGGATGAGCTCAAGCTGCGGCCAAGCT 2619
QY	1621	GAGTGTATCAAGGATCTCTGGCGCGCTGCGGAGTCTCTCTCAATTCATCTCTCTCAAGAT 1680
DB	2620	GAGTGTATCAAGGATCTCTGGCGCGCTGCGGAGTCTCTCTCAATTCATCTCTCTCAAGAT 2679
QY	1681	CACCTCGAGAAAGTCAAGGCACTTCGAGAGGAAATTCGCGCTCTGAAAGAGAACTGAGC 1740
DB	2680	CACCTCGAGAAAGTCAAGGCACTTCGAGAGGAAATTCGCGCTCTGAAAGAGAACTGAGC 2739
QY	1741	CAGGTCAATGACCTTGTCTGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAC 1800
DB	2740	CAGGTCAATGACCTTGTCTGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAC 2799

1801 CTCACACTCTGGAGAGCTGAACACACAGATGGAAGCTTCTGAGGTGGCGTCGAGGAC 1860
1861 CGAGTCAGGAGCTGCATGAAGCCACACAGGAGCTTTGGTCCAGCATCTCAGCATTCTTT 1920
1921 TCCAGTCTGTCCAGGCTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTCCCTACTAT 1980
1981 ATCAACACAGAGACTCAAAACA 2001
2001 ATCAACACAGAGACTCAAAACA 3000
2002 ATCAACACAGAGACTCAAAACA 3000
RESULT 2
AD37256
D AAD37256 standard; DNA; 4966 BP.
X C AAD37256;
X T 21-AUG-2002 (first entry)
X Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
X Human; dystrophin minigene; muscular; Gene therapy; utrophin; spectrin;
W adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
W Becker muscular dystrophy; ds.
X Chimeric - Homo sapiens.
S Chimeric - Unidentified.
S WO200183695-A2.
X 08-NOV-2001.
X 27-APR-2001; 2001WO-US13677.
X 28-APR-2000; 2000US-200777P.
X (XIAO/) XIAO X.
X Xiao X;
X WPI; 2002-049342/06.
X New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
X Example 1; Page 59-60; 71pp; English.
X The present invention relates to an isolated nucleotide sequence encoding
X a dystrophin minigene. The minigene comprises N-terminal or modified
X N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
X domains and cysteine-rich domains of dystrophin or utrophin genes. The
X invention also relates to a recombinant adeno-associated virus (AAV)
X comprising dystrophin minigene operably linked to an expression control
X element. The dystrophin minigene in operable linkage with an expression
X control element, in a recombinant adeno-associated virus or retrovirus is
X useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
X dystrophy (BMD) in a mammalian subject. The present sequence is AAV
X vector plasmid construct containing human dystrophin minigenes, a muscle
X creatine kinase (CK) promoter and a small polyA signal sequence.
X
X Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;
Query March 100.0%; Score 2001; DB 24; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTTTCATTCGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
DB 1757 GGCAGTTTCATTCGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1816
QY 61 GTATTATCGTGGCTCTTTCTCTGCTGAGGACACATTTGCAAGCAAAAGGAGAGATTTCTTAT 120
DB 1817 GTATTATCGTGGCTCTTTCTCTGCTGAGGACACATTTGCAAGCAAAAGGAGAGATTTCTTAT 1876
QY 121 GATGTGGAAGTGGTGAAGAGACAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 180
DB 1877 GATGTGGAAGTGGTGAAGAGACAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 1936
QY 181 GCCCATCAGGCGGGTGGTAAATTTCTCAAAATGGGAAGTAAAGTCTGATTTGGAACAGGA 240
DB 1937 GCCCATCAGGCGGGTGGTAAATTTCTCAAAATGGGAAGTAAAGTCTGATTTGGAACAGGA 1996
QY 241 AAATTATCAGAAGTGAAGAACTGAAGTACAGAGCAGATGAATCTCTCAAAATCAAGA 300
DB 1997 AAATTATCAGAAGTGAAGAACTGAAGTACAGAGCAGATGAATCTCTCAAAATCAAGA 2056
QY 301 TGGGAATGCTCTCAGGCTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAAATG 360
DB 2057 TGGGAATGCTCTCAGGCTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAAATG 2116
QY 361 GATCTCAGATCAGAACTGAAGAGTTGAATGCTGCTAAGAAACAGAAAGAAAGA 420
DB 2117 GATCTCAGATCAGAACTGAAGAGTTGAATGCTGCTAAGAAACAGAAAGAAAGA 2176
QY 421 ACAAGGAAATGGAGGAGAGGCTCTTTGAGCCTGATCTTTGAAGAGCTTAAACGCCAAGTA 480
DB 2177 ACAAGGAAATGGAGGAGAGGCTCTTTGAGCCTGATCTTTGAAGAGCTTAAACGCCAAGTA 2236
QY 481 CAACAACATPAAGTGTCTTCAAGAGATCTAGAAACAGAAACAGTCAAGGTCAATTTCTCTC 540
DB 2237 CAACAACATPAAGTGTCTTCAAGAGATCTAGAAACAGAAACAGTCAAGGTCAATTTCTCTC 2296
QY 541 ACTCACATGCTGGTGTAGTTGATGAATCTAGTGAGATCAGGCAACTGCTGCTTTGGAA 600
DB 2297 ACTCACATGCTGGTGTAGTTGATGAATCTAGTGAGATCAGGCAACTGCTGCTTTGGAA 2356
QY 601 GAACAACATTAAGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 560
DB 2357 GAACAACATTAAGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCCGC 2416
QY 561 TGGGTTCTTTTACAAGACCAAGCTGACCTAGCTCTGAGTACCACTATTGGAGCTCT 720
DB 2417 TGGGTTCTTTTACAAGACCAAGCTGACCTAGCTCTGAGTACCACTATTGGAGCTCT 2476
QY 721 CCTACTCAGACTGTACTCTCTGGTGACACAACTGTGGTTTACTAAGGAACTGCCATCTCC 780
DB 2477 CCTACTCAGACTGTACTCTCTGGTGACACAACTGTGGTTTACTAAGGAACTGCCATCTCC 2536
QY 781 AAATAGAAATGCACTCTCTCTGAGTGGAGGTACTTACTCATAGATTACTGCAACAG 840
DB 2537 AAATAGAAATGCACTCTCTCTGAGTGGAGGTACTTACTCATAGATTACTGCAACAG 2596
QY 841 TTCCCTCTGACCTGGAAAAGTTTCTTGCTTGGCTTACAGAAGCTGAAAACAACTGCCAAT 900
DB 2597 TTCCCTCTGACCTGGAAAAGTTTCTTGCTTGGCTTACAGAAGCTGAAAACAACTGCCAAT 2656
QY 901 GTCTACAGAGTGTACCTCCCTAGGAAAGCTCTCTAGAGACTCCCAAGGGAGTAAAGAG 960
DB 2657 GTCTACAGAGTGTACCTCCCTAGGAAAGCTCTCTAGAGACTCCCAAGGGAGTAAAGAG 2716
QY 961 CTGATGAAACAAATGGCAAGACCTCCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 1020
DB 2717 CTGATGAAACAAATGGCAAGACCTCCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAC 2776
QY 1021 AACCTGATGAAACAGCCAAATAATCTCTGAGATCCCTGGAAGTTCCTGATGATGAGTTC 1080
DB 2777 AACCTGATGAAACAGCCAAATAATCTCTGAGATCCCTGGAAGTTCCTGATGATGAGTTC 2836

QY	1081	CTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCT	1140
DB	2837	CTGTACAAAGACGTTTGGATAAACAATGAACCTTCAAGTGGAGTGAACCTTCGGAAAAAGTCT	2896
QY	1141	CTCAACATTAGTTCACATTGGNAGCCAGTTCTGACCACTGGAAAGCGTCTGCACCTTTCT	1200
DB	2897	CTCAACATTAGTTCACATTGGNAGCCAGTTCTGACCACTGGAAAGCGTCTGCACCTTTCT	2956
QY	1201	CTCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACT	1260
DB	2957	CTCAGGAACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACT	3016
QY	1261	ATTGGAGGCGACTTTCCAGCAGTTTCAGACGAGCAAGCGATGTACATAGGSCCTTCAAGAGG	1320
DB	3017	ATTGGAGGCGACTTTCCAGCAGTTTCAGACGAGCAAGCGATGTACATAGGSCCTTCAAGAGG	3076
QY	1321	GAATTGAATACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTCTACGAATATTTCG	1380
DB	3077	GAATTGAATACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTCTAGCAATATTTCG	3136
QY	1381	ACAGACGAGCCTTTGGNAGCACTAGAGAACTCTACCGAGGCCCCAGAGAGCTGCCTCT	1440
DB	3137	ACAGACGAGCCTTTGGNAGCACTAGAGAACTCTACCGAGGCCCCAGAGAGCTGCCTCT	3196
QY	1441	GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGCTGAGAGSTCAATACT	1500
DB	3197	GAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGCTGAGAGSTCAATACT	3256
QY	1501	GAGTGGGAAAAATTGAACCTGCACTCGCTGACTGCGCAGAGAAATAGATGAGACCCCT	1560
DB	3257	GAGTGGGAAAAATTGAACCTGCACTCGCTGACTGCGCAGAGAAATAGATGAGACCCCT	3316
QY	1561	GAAGAATCTCAGGAACTTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCT	1620
DB	3317	GAAGAATCTCAGGAACTTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCCAAAGCT	3376
QY	1621	GAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGATCTCTCTCATGTGACTCTCTCCAGAT	1680
DB	3377	GAGGTGATCAAGGGATCTTGGCAGCCCGTGGCGATCTCTCTCATGTGACTCTCTCCAGAT	3436
QY	1681	CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTTGAAAGAGAACGTGAGC	1740
DB	3437	CACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTGCGCTCTTGAAAGAGAACGTGAGC	3496
QY	1741	CAGTCAATGACCTTGTCTGCCACTTACCACTTTGGGCATTGAGCTCTCACCGTATAAC	1800
DB	3497	CAGTCAATGACCTTGTCTGCCACTTACCACTTTGGGCATTGAGCTCTCACCGTATAAC	3556
QY	1801	CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGAC	1860
DB	3557	CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGAC	3616
QY	1861	CGAGTCAGGCAAGCTGATGAAGCCCAAGGACCTTGGTCAGGATCTCAGCACTTTCTT	1920
DB	3617	CGAGTCAGGCAAGCTGATGAAGCCCAAGGACCTTGGTCAGGATCTCAGCACTTTCTT	3676
QY	1921	TCCAGCTCTGTCCAGGTCCTTGGAGAGAGCCATCTCGCCAAAACAAAGTCCCTACTAT	1980
DB	3677	TCCAGCTCTGTCCAGGTCCTTGGAGAGAGCCATCTCGCCAAAACAAAGTCCCTACTAT	3736
QY	1981	ATCAACACGAGACTCAAAACA	2001
DB	3737	ATCAACACGAGACTCAAAACA	3757

RESULT 3
AAD37262

XX AAD37262:

XX
DT 21-AUG-2002 (first entry)

XXII

DE	Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
XX	Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX	adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX	Becker muscular dystrophy; ds.
XX	Chimeric - Homo sapiens.
OS	Chimeric - Cytomegalovirus.
OS	Chimeric - Unidentified.
XX	WO200183695-A2.
PN	08-NOV-2001.
XX	27-APR-2001; 2001WO-US13677.
XX	28-APR-2000; 2000US-200777P.
XX	(XIAO/) XIAO X.
XX	Xiao X;
XX	WPI; 2002-049342/06.
DR	New dystrophin minigene for treating Duchenne or Becker muscular
XX	dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT	rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT	dystrophin gene -
PT	

2

rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -

Example 1; Page 48-49; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus, hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;

Query Match	85.4%;	Score 1709;	DB 24;	Length 3858;
Best Local Similarity	93.0%;	Pred. No. 0;		
Matches 1860;	Conservative 0;	Mismatches 0;	Indels 141;	Gaps 1;
1	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGAGCTTTAGAAGAA	60		
1000	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGAGCTTTAGAAGAA	1059		
61	GTATTATGCTGGCTTCTCTGCTGAGACACATTCGAAGCACAAGGAGAGATTTCTAAT	120		
1060	GTATTATGCTGGCTTCTCTGCTGAGACACATTCGAAGCACAAGGAGAGATTTCTAAT	1119		
121	GATGTGGAAAGTGGTGAAGACCCAGTTCATCTCATGAGGGGTATCATGATGATTTGACA	180		
1120	GATGTGGAAAGTGGTGAAGACCCAGTTCATCTCATGAGGGGTATCATGATGATTTGACA	1179		
181	GCCATCAGGGCGGGTGGTAAATTTCTAATTTGGAAGTGAAGTGAAGTGAAGTGAAGT	240		
1180	GCCATCAGGGCGGGTGGTAAATTTCTAATTTGGAAGTGAAGTGAAGTGAAGTGAAGT	1239		
241	AAATTATCAGAGATGAAGAACTCAAGTACAGAGCAGATGAATCTCTCTAAATTCAGA	300		
1240	AAATTATCAGAGATGAAGAACTCAAGTACAGAGCAGATGAATCTCTCTAAATTCAGA	1299		
301	TGGAAATGCTCAGGGTAGCTAGCATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	360		
1300	TGGAAATGCTCAGGGTAGCTAGCATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	1359		
361	GATCTCCGAATCAGAACTGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	420		
1360	GATCTCCGAATCAGAACTGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	1419		
421	ACAAGGAAATGGAGAGAGGCTCTTGACCTGATCTTGAAGAGCTTAAACGCCAAGTA	480		
1420	ACAAGGAAATGGAGAGAGGCTCTTGACCTGATCTTGAAGAGCTTAAACGCCAAGTA	1479		
481	CAACAACATGAAGTCTCAAGAGATCTAGAACAGAGCAAGTCAAGGTCAATTTCTC	540		
1480	CAACAACATGAAGTCTCAAGAGATCTAGAACAGAGCAAGTCAAGGTCAATTTCTC	1539		
541	ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	600		
1540	ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1599		
601	GACCACTTAAGTATTTGGAGATCATGGGCAACATCTGTAGATGGACAGACCCG	660		
1600	GACCACTTAAGTATTTGGAGATCATGGGCAACATCTGTAGATGGACAGACCCG	1659		
661	TGGTCTTTTACAAGACACAGCCTAGCTCTCTGGAGTCACTGACCACTATTGGAGCTCT	720		
1660	TGGTCTTTTACAAGACACAGCCTAGCTCTCTGGAGTCACTGACCACTATTGGAGCTCT	1677		

QY	721	CCCTACTCAGACTGTTTACTCTGGTGACACAACTGTGGTTACTAAGGAAACTGGCCATCTCC	780
DB	1678	-----	1677
QY	781	AAACTAGAAATGCCATCTCTCTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAG	840
DB	1678	-----	1678
QY	841	TTCCCCCTGGACCTGGAAAAGTTTCTTCCCTGGTTACAGAAGCTGAAAACACTGCCAAT	900
DB	1699	TTCCCCCTGGACCTGGAAAAGTTTCTTCCCTGGTTACAGAAGCTGAAAACACTGCCAAT	1758
QY	901	GTCTTACAGATGCTACCGGTAGAAAGGCTCTCTAGAACTCCAAAGGAGTAAACAG	960
DB	1759	GTCTTACAGATGCTACCGGTAGAAAGGCTCTCTAGAACTCCAAAGGAGTAAACAG	1818
QY	961	CTGTGAAACAAATGCGAAGACCTCCAAAGGTGAAATTTGAAGCTCACAGATGTTTATCAC	1020
DB	1819	CTGTGAAACAAATGCGAAGACCTCCAAAGGTGAAATTTGAAGCTCACAGATGTTTATCAC	1878
QY	1021	AACTGGATGAAACAGCCAAAATAATCTCTGATCCCTGGAGGTTCCGATGATGATGATC	1080
DB	1879	AACTGGATGAAACAGCCAAAATAATCTCTGATCCCTGGAGGTTCCGATGATGATGATC	1938
QY	1081	CTGTTACAAAGACGTTTGGATAAATCAATGAATTTCAAGTGGAGTGAATTTGGAAAAGTCT	1140
DB	1939	CTGTTACAAAGACGTTTGGATAAATCAATGAATTTCAAGTGGAGTGAATTTGGAAAAGTCT	1998
QY	1141	CTCAACATTTAGTCCCATTTTGAAGCCAGTCTTGACAGTGGAAAGCTCTGCACCTTTCT	1200
DB	1999	CTCAACATTTAGTCCCATTTTGAAGCCAGTCTTGACAGTGGAAAGCTCTGCACCTTTCT	2058
QY	1201	CTGCAAGAACTTTCTGTTGGCTTACAGCTGAAAGATGATGAATTTAAGCCGCGAGGCACT	1260
DB	2059	CTGCAAGAACTTTCTGTTGGCTTACAGCTGAAAGATGATGAATTTAAGCCGCGAGGCACT	2118
QY	1261	ATTGGAGGCGACTTTCCAGCAGTTTCAAGACGAGTATGATGATGATGATGATGATGATG	1320
DB	2119	ATTGGAGGCGACTTTCCAGCAGTTTCAAGACGAGTATGATGATGATGATGATGATGATG	2178
QY	1321	GAATTTGAAACTTAAAGAACTCTGATGATGATGATGATGATGATGATGATGATGATGATG	1380
DB	2179	GAATTTGAAACTTAAAGAACTCTGATGATGATGATGATGATGATGATGATGATGATGATG	2238
QY	1381	ACAGAGCAGCCTTTTGAAGGAGCTTAGAGAACTCTACAGAGCCCGAGAGCTGCTCTCT	1440
DB	2239	ACAGAGCAGCCTTTTGAAGGAGCTTAGAGAACTCTACAGAGCCCGAGAGCTGCTCTCT	2298
QY	1441	GAGGAGAGCCCGAGAGTCTCTGCTGCTTCTACGAAAGCAGGCTGAGGAGTCAATCT	1500
DB	2299	GAGGAGAGCCCGAGAGTCTCTGCTGCTTCTACGAAAGCAGGCTGAGGAGTCAATCT	2358
QY	1501	GAGTGGGAAAATTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1560
DB	2359	GAGTGGGAAAATTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2418
QY	1561	GAAAGACTCCAGAACTTCAAGAGCCAGATGATGATGATGATGATGATGATGATGATGATG	1620
DB	2419	GAAAGACTCCAGAACTTCAAGAGCCAGATGATGATGATGATGATGATGATGATGATGATG	2478
QY	1621	GAGGTGATCAAGGATCTCTGGCAGCCCTGGGGGATCTCTCTCATTTCACTCTCTCTC	1680
DB	2479	GAGGTGATCAAGGATCTCTGGCAGCCCTGGGGGATCTCTCTCATTTCACTCTCTCTC	2538
QY	1681	CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCGCTCTGAAAGAGACGTTGAG	1740
DB	2539	CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCGCTCTGAAAGAGACGTTGAG	2598
QY	1741	CAGTCAATGACCTTCTGCTCCAGCTTACCACTTTTGGGATTCAGTCTCTCACCGTATAAC	1800
DB	2599	CAGTCAATGACCTTCTGCTCCAGCTTACCACTTTTGGGATTCAGTCTCTCACCGTATAAC	2658
QY	1801	CTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCGAGTGGCGCTCGAGGAC	1860

Db 2696 CTGTTACAGACGCTTTGGATAACATGACTTCAAGTGGAGTGAACCTCGGAAAAAGTCT 2755
Qy 1141 CTCACATTTAGTCCCAATTTGGAAAGCAGTCTGACAGTGGAGCGTCTGACCTTTCT 1200
Db 2756 CTCACATTTAGTCCCAATTTGGAAAGCAGTCTGACAGTGGAGCGTCTGACCTTTCT 2815
Qy 1201 CTGAGGAACCTTCGTGGTGTGCTACAGTGAAGATGATGAATTAAGCCGGCAGGACCT 1260
Db 2816 CTGAGGAACCTTCGTGGTGTGCTACAGTGAAGATGATGAATTAAGCCGGCAGGACCT 2875
Qy 1261 ATTGGAGCGACTTTCAGCAGTTCAGAGCAGAACGATGATACATAGGGCTTCAGAGG 1320
Db 2876 ATTGGAGCGACTTTCAGCAGTTCAGAGCAGAACGATGATACATAGGGCTTCAGAGG 2935
Qy 1321 GAAATTGAAACTTAAGAACCTTGAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 1380
Db 2936 GAAATTGAAACTTAAGAACCTTGAATCATGAGTACTCTTGAGACTGTACGAATATTTCTG 2995
Qy 1381 ACAGAGCAGCCTTTGGAAAGGACTAGAGAACTCTACAGGAGCCAGAGTGCCTCCT 1440
Db 2996 ACAGAGCAGCCTTTGGAAAGGACTAGAGAACTCTACAGGAGCCAGAGTGCCTCCT 3055
Qy 1441 GAGGAGAGCCAGAACTGCTACTCGCTTCTAGAGAGCAGGCTGAGAGTCAATACT 1500
Db 3056 GAGGAGAGCCAGAACTGCTACTCGCTTCTAGAGAGCAGGCTGAGAGTCAATACT 3115
Qy 1501 GAGTGGGAAAAATTAAGACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTT 1560
Db 3116 GAGTGGGAAAAATTAAGACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTT 3175
Qy 1561 GAAAGACTCCAGAACTTCAAGAGCCAGGATGAGTGGACCTCAAGCTCGCCAGCT 1620
Db 3176 GAAAGACTCCAGAACTTCAAGAGCCAGGATGAGTGGACCTCAAGCTCGCCAGCT 3235
Qy 1621 GAGGTGATCAAGGATCTCTGGCAGCCCTGGCGGATCTCTCATGACTCTCTCAAGAT 1680
Db 3236 GAGGTGATCAAGGATCTCTGGCAGCCCTGGCGGATCTCTCATGACTCTCTCAAGAT 3295
Qy 1681 CACCTCGAAGAGTCAAGCATTGAGAGAAATGGCCCTCTGAAGAGAGAGTGAAGC 1740
Db 3296 CACCTCGAAGAGTCAAGCATTGAGAGAAATGGCCCTCTGAAGAGAGAGTGAAGC 3355
Qy 1741 CAGCTCAATGACCTTCTCGCAGCTTACACTTTGGGCATTCAGCTCTCAACCGTATTAAC 1800
Db 3356 CAGCTCAATGACCTTCTCGCAGCTTACACTTTGGGCATTCAGCTCTCAACCGTATTAAC 3415
Qy 1801 CTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCCCTCAGGAC 1860
Db 3416 CTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCCCTCAGGAC 3475
Qy 1861 CGAGTCAGGAGCTGATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTT 1920
Db 3476 CGAGTCAGGAGCTGATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTT 3535
Qy 1921 TCCAGCTCTCTCCAGGCTCTCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTAT 1980
Db 3536 TCCAGCTCTCTCCAGGCTCTCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTAT 3595
Qy 1981 ATCAACACAGAGCTCAACA 2001
Db 3596 ATCAACACAGAGCTCAACA 3616

RESULT 6
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
XX
AC
XX
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13677.
XX
XX 28-APR-2000; 2000US-200777P.
XX
XX (XIAO/) XIAO X.
XX
XX Xiao X;
XX
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene
XX
XX Example 1; Page 68-70; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence.
XX
XX Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;
XX
XX Query Match 85.4%; Score 1709; DB 24; Length 4848;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 0; Indels 141; Gaps 1;
Qy 1 GGCAGTTCATTGATGGAGTGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAGAA 60
Db 1780 GGCAGTTCATTGATGGAGTGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAAGAA 1839
Qy 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 120
Db 1840 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 1899
Qy 121 GATGTGAAGTGTGAAGACCACTGTTACTCATGAGGGGTACATGATGGATTGACA 180
Db 1900 GATGTGAAGTGTGAAGACCACTGTTACTCATGAGGGGTACATGATGGATTGACA 1959
Qy 181 GCCCATCAGGGCCGGGTGGTGAATATTCTCAATTTGGGAAGTAAAGCTGATTGGAACAGGA 240
Db 1960 GCCCATCAGGGCCGGGTGGTGAATATTCTCAATTTGGGAAGTAAAGCTGATTGGAACAGGA 2019
Qy 241 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGATCTCTTAAATTTCAAGA 300
Db 2020 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGATCTCTTAAATTTCAAGA 2079
Qy 301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTAATG 360
Db 2080 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGATTTAATG 2139

361 GATCTCCAGAACTCAGAACTGAAAGAGTTGAATGACTGCTAAACAAACAGAGAAAGA 420
2140 GATCTCCAGAACTCAGAACTGAAAGAGTTGAATGACTGCTAAACAAACAGAGAAAGA 2199
421 ACAAGAAAAATGGAGAGAGAGCTCTTGGAAGCTGATCTTTGAAGAGCTAAACAGCCAGTA 480
2200 ACAAGAAAAATGGAGAGAGAGCTCTTGGAAGCTGATCTTTGAAGAGCTAAACAGCCAGTA 2259
481 CAACACATAAGGTGCTTCAAGAGAGATCTAGAAACAGAACTAGTCAAGGTCAATTTCTCTC 540
2260 CAACACATAAGGTGCTTCAAGAGAGATCTAGAAACAGAACTAGTCAAGGTCAATTTCTCTC 2319
541 ACTCACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
2320 ACTCACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2379
601 GAACAACTTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
2380 GAACAACTTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2439
661 TGGGTTCTTTTACAAACACAGAGCTGACCTAGCTCTCTGAGTCAACCACTATTGAGGCTCT 720
2440 TGGGTTCTTTTACAAACACAGAGCTGACCTAGCTCTCTGAGTCAACCACTATTGAGGCTCT 2457
721 CCTACTCAGACTGTTACTCTGTTGACACAACTGTTGTTACTTAAGGAACACTGCCATCTCC 780
2458 ----- 2457
781 AAACCTAGAAATGCCATCTTCTCTGATGTTGGAGGTACTTACTATGATGATGATGATGAT 840
2458 -----ACTCATAGATTACTGCAACAG 2478
841 TTCCCTGCTGACCTGAAAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
2479 TTCCCTGCTGACCTGAAAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2538
901 GTCTCTCAGAGATGCTACCGCTAGAGAGAGCTCCTTAGAAGACTCCAAAGGAGTAAAGAG 960
2539 GTCTCTCAGAGATGCTACCGCTAGAGAGAGCTCCTTAGAAGACTCCAAAGGAGTAAAGAG 2598
961 CTGATGAAACAAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACAAGATGTTATCAC 1020
2599 CTGATGAAACAAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACAAGATGTTATCAC 2658
1021 AACCTGATGAAACAAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACAAGATGATGCTC 1080
2659 AACCTGATGAAACAAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACAAGATGATGCTC 2718
1081 CTGTTACAAAGACCTTTGGGATAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCT 1140
2719 CTGTTACAAAGACCTTTGGGATAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCT 2778
1141 CTCAACTTAGTCCCAATTTGGAGAGCAGTCTGACAGTGGAGGCTGTCACCTTTCT 1200
2779 CTCAACTTAGTCCCAATTTGGAGAGCAGTCTGACAGTGGAGGCTGTCACCTTTCT 2838
1201 CTGAGAGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
2839 CTGAGAGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2898
1261 ATTGGAGGCGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTTCAAGAGG 1320
2899 ATTGGAGGCGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTTCAAGAGG 2958
1321 GAATTTGAAGAACTTAAGACCTGTAATCATGAGTACTCTTGGAGTGTACGAATATTCTG 1380
2959 GAATTTGAAGAACTTAAGACCTGTAATCATGAGTACTCTTGGAGTGTACGAATATTCTG 3018
1381 ACAGAGCAGCTTTTGGAGGAGTACAGAACTTACAGAGAGCCAGAGAGCTGCTCTCT 1440
3019 ACAGAGCAGCTTTTGGAGGAGTACAGAACTTACAGAGAGCCAGAGAGCTGCTCTCT 3078
1441 GAGGAGAGAGCCAGAGATGCTACCTGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACT 1500

Db 3079 GAGGAGAGAGCCAGAAATGTCACCTGGCTTTACGAAACAGGCTGAGAGGTCAATACT 3138
QY 1501 GAGTGGGAAAAATTTGAACCTGCACTCCGCTGCACTGGCAGAGAAAAATAGATGAGACCTT 1560
Db 3139 GAGTGGGAAAAATTTGAACCTGCACTCCGCTGCACTGGCAGAGAAAAATAGATGAGACCTT 3198
QY 1561 GAAAGACTCCAGAAATTTCAAGAGGCCACGAGTGAAGCTCAAGCTGGGCCAAGCT 1620
Db 3199 GAAAGACTCCAGAAATTTCAAGAGGCCACGAGTGAAGCTCAAGCTGGGCCAAGCT 3258
QY 1621 GAGGTGATCAAGGATCTCTGGAGGCCCTGGAGGCCGATCTCTCAATTCAGCTCTCTCCCAAGAT 1680
Db 3259 GAGGTGATCAAGGATCTCTGGAGGCCCTGGAGGCCGATCTCTCAATTCAGCTCTCTCCCAAGAT 3318
QY 1681 CACCTCGAGAAATCAAGGCACTTCGAGAGGAAATTTGGCTCTGAAAGAGAACTGAGC 1740
Db 3319 CACCTCGAGAAATCAAGGCACTTCGAGAGGAAATTTGGCTCTGAAAGAGAACTGAGC 3378
QY 1741 CAGTCAATGACCTTCTGCTCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC 1800
Db 3379 CAGTCAATGACCTTCTGCTCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAC 3438
QY 1801 CTGAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGTGGCCGCTCGAGGAC 1860
Db 3439 CTGAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTGAGTGGCCGCTCGAGGAC 3498
QY 1861 CGAGTCAGGAGCTGCAATGAAGCCACAGGACTTTGGTCCAGCACTCTCAGCACTTTCTT 1920
Db 3499 CGAGTCAGGAGCTGCAATGAAGCCACAGGACTTTGGTCCAGCACTCTCAGCACTTTCTT 3558
QY 1921 TCCAGCTCTGTCAGGCTCTCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 1980
Db 3559 TCCAGCTCTGTCAGGCTCTCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTAT 3618
QY 1981 ATCAACCAAGAGACTCAACA 2001
Db 3619 ATCAACCAAGAGACTCAACA 3639

RESULT 7
AAD37264
ID AAD37264 standard; DNA; 5060 BP.
XX
AC AAD37264;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
XX
XX Chimeric - Unidentified.
XX
XX WO200183695-A2.
XX
PD 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13677.
XX
XX 28-APR-2000; 2000US-200777P.
XX
XX (XIAO/) XIAO X.
XX
XX Xiao X;
XX
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular

1801 CTGAGCACTCTGGAAGACCTGAACACAGATGAAGCTTCTGAGGTGGCCCTGAGGAC 1860
3651 CTGAGCACTCTGGAAGACCTGAACACAGATGAAGCTTCTGAGGTGGCCCTGAGGAC 3710
1861 CGAGTCAGGAGCTGATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGACATTTCTT 1920
3711 CGAGTCAGGAGCTGATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGACATTTCTT 3770
1921 TCCACGTCTGTCCAGGCTCCCTGGAGAGAGCCATCTCCGCAAAACAAAGTGCCTACTAT 1980
3771 TCCACGTCTGTCCAGGCTCCCTGGAGAGAGCCATCTCCGCAAAACAAAGTGCCTACTAT 3830
1981 ATCAACACGAGACTCAACA 2001
3831 ATCAACACGAGACTCAACA 3851

RESULT 8

AAD37230

ID AAD37230 standard; DNA; 4182 BP.

XX

AC

AAD37230;

DT 21-AUG-2002 (first entry)

XX

Human dystrophin minigene delta4173.

XX

Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX

Homo sapiens.

XX

WO200183695-A2.

XX

08-NOV-2001.

XX

27-APR-2001; 2001WO-US13677.

XX

28-APR-2000; 2000US-200777P.

XX

(XIAO/) XIAO X.

XX

Xiao X;

XX

WPI; 2002-049342/06.

XX

New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene

XX

Example 1; Page 43-44; 71pp; English.

XX

The present invention relates to an isolated nucleotide sequence encoding

CC a dystrophin minigene. The minigene comprises N-terminal or modified

CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

CC domains and cysteine-rich domains of dystrophin or utrophin genes. The

CC invention also relates to a recombinant adeno-associated virus (AAV)

CC comprising dystrophin minigene operably linked to an expression control

CC element. The dystrophin minigene in operable linkage with an expression

CC control element, in a recombinant adeno-associated virus or retrovirus is

CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC dystrophy (BMD) in a mammalian subject. The present sequence is human

CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,

CC hinge H1 and rod repeats R1, R2 and R3), 8059-10227 (rod repeats R2, R3 and R4,

CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX

Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 other;

XX

Query Match 83.1%; Score 1662.8; DB 24; Length 4182;

XX

Best Local Similarity 88.3%; Pred. No. 0;

XX

Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

QY 1 GGCAATTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 60
DB 1000 GGCAATTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 1059
QY 61 GTATTATCGTGGCTCTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAAT 120
DB 1060 GTATTATCGTGGCTCTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAAT 1119
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCTACTCATGAGGGGTACATGATGGATTTGACA 180
DB 1120 GATGTGGAAGTGGTGAAGACCAAGTTTCTACTCATGAGGGGTACATGATGGATTTGACA 1179
QY 181 GCCCATCAGGCGCGGTGGTGAATATTTCTCAATTTGGGAAGTAAAGCTGATGGAACAGGA 240
DB 1180 GCCCATCAGGCGCGGTGGTGAATATTTCTCAATTTGGGAAGTAAAGCTGATGGAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAAACCTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 300
DB 1240 AAATTATCAGAAGATGAAGAAACCTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 1299
QY 301 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAGAAACCAAGCAATTTACATAGATTTTAATG 360
DB 1300 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAGAAACCAAGCAATTTACATAGATTTTAATG 1359
QY 361 GATCTCCAGATCAGAACTGAAGAGTTGAATGCTTAACTGCTTAAACCAAGCAAGGAAGA 420
DB 1360 GATCTCCAGATCAGAACTGAAGAGTTGAATGCTTAACTGCTTAAACCAAGCAAGGAAGA 1419
QY 421 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTA 480
DB 1420 ACAAGGAAATGGAGGAAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTA 1479
QY 481 CAACAACATAGGTGCTTCAAGAGATCTAGACAGACCAAGTCAAGGTCAATTTCTCTC 540
DB 1480 CAACAACATAGGTGCTTCAAGAGATCTAGACAGACCAAGTCAAGGTCAATTTCTCTC 1539
QY 541 ACTCAGATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 600
DB 1540 ACTCAGATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1599
QY 601 GAACAACATTAAGTATGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGC 660
DB 1600 GAACAACATTAAGTATGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGC 1659
QY 661 TGGGTCTCTTTTACAAGACCAAGCCTGAC-----CTAGCTCTCTGGACTGACCACT 708
DB 1660 TGGGTCTCTTTTACAAGACATCTTCTCAATGGCAACGCTCTTACTGGAAGACAGTGCCTT 1719
QY 709 ATTGGAGCCTCTCTCTACTCAGACTGTTTACTCTGGTGACACA-----ACCTGTGGTT 759
DB 1720 TTTAGTGCATGGCTTTTCAAGAAAGAGATGCGATGGGCAACATCTGTAGATGGACAGAGACCGC 1779
QY 760 ACTAAGGAACTGCCATCTC----- 779
DB 1780 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTAAAGCGGATCTA 1839
QY 780 -----CAAACCTAGAAATGCCATCTTCC 801
DB 1840 GAAAGAAAGCAATCCATGGCAAACTGTATTCTCACTCAACCAAGATCTTCTTTCAACA 1899
QY 802 TTGATGTTGGAG----- 813
DB 1900 CTGAAGATAAGTCAAGTGAACCCAGAGACCGGATGCGGTGGATTAATCTTGGCCGGTGT 1959
QY 814 -----GTACCTACTCATAGATTTACTGCAA 837
DB 1960 TGGGATAATTTAGTCCAAAACCTTGAAGAGAGTACAGACAGACTCATAGATTTACTGCAA 2019
QY 838 CAGTTTCCCTCGGACCTGGAAAGTTTCTGCTGCTGGCTTACAGAGCTGAACAACTGCC 897
DB 2020 CAGTTTCCCTCGGACCTGGAAAGTTTCTGCTGCTGGCTTACAGAGCTGAACAACTGCC 2079

QY 898 AATGTCCTACAGGATGCTACCCGTAAGAAAGGCTCTAGAGACTCCAGGGAGTAAA 957
Db |||||
QY 2080 AATGTCCTACAGGATGCTACCCGTAAGAAAGGCTCTAGAGACTCCAGGGAGTAAA 2139
Db |||||
QY 958 GAGCTGATGAACCAATGCAAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTAT 1017
Db |||||
QY 2140 GAGCTGATGAACCAATGCAAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTAT 2199
Db |||||
QY 1018 CACACCTGGATGAAAAAGCAGCCAAAAATCTGAGATCCCTGGAGGTTCCGATGATGCA 1077
Db |||||
QY 2200 CACACCTGGATGAAAAAGCAGCCAAAAATCTGAGATCCCTGGAGGTTCCGATGATGCA 2259
Db |||||
QY 1078 GTCCTGTTACAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAATTTGCGAAAAAG 1137
Db |||||
QY 2260 GTCCTGTTACAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAATTTGCGAAAAAG 2319
Db |||||
QY 1138 TCTCTCAACATTAGTCCATTGGAAGCAGCTTCTGACAGTGGAGGCTCTGCACCTT 1197
Db |||||
QY 2320 TCTCTCAACATTAGTCCATTGGAAGCAGCTTCTGACAGTGGAGGCTCTGCACCTT 2379
Db |||||
QY 1198 TCTCTGCGAGACTTCTGTGTGGCTTACAGCTGAAAGATGATGAATTAAGCCGCGAGCA 1257
Db |||||
QY 2380 TCTCTGCGAGACTTCTGTGTGGCTTACAGCTGAAAGATGATGAATTAAGCCGCGAGCA 2439
Db |||||
QY 1258 CCTATTGGAGGCGACTTCCACAGCTTCAAGAGCAGATGATACATAGGCGCTTCAAG 1317
Db |||||
QY 2440 CCTATTGGAGGCGACTTCCACAGCTTCAAGAGCAGATGATACATAGGCGCTTCAAG 2499
Db |||||
QY 1318 AGGGAATTGAAACTTAAGAACTGTAATCATGACTCTTGTGACGTGATGATATTT 1377
Db |||||
QY 2500 AGGGAATTGAAACTTAAGAACTGTAATCATGACTCTTGTGACGTGATGATATTT 2559
Db |||||
QY 1378 CTGACAGAGCAGCTTTTGAAGGACTTAGAGAACTCTACAGGAGCCCGAGAGCTGCT 1437
Db |||||
QY 2560 CTGACAGAGCAGCTTTTGAAGGACTTAGAGAACTCTACAGGAGCCCGAGAGCTGCT 2619
Db |||||
QY 1438 CTTGAGGAGAGAGCCAGATGCTACTGGCTTCTACGAGAGAGGCTGAGAGGTCAT 1497
Db |||||
QY 2620 CTTGAGGAGAGAGCCAGATGCTACTGGCTTCTACGAGAGAGGCTGAGAGGTCAT 2679
Db |||||
QY 1498 ACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGCGCAGAGAAAAATAGTAGAGC 1557
Db |||||
QY 2680 ACTGAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGCGCAGAGAAAAATAGTAGAGC 2739
Db |||||
QY 1558 CTTGAAAGACTCCAGAACTTCAAGAGGCCAGGATGAGCTGAGCTCAAGTGGCCAA 1617
Db |||||
QY 2740 CTTGAAAGACTCCAGAACTTCAAGAGGCCAGGATGAGCTGAGCTCAAGTGGCCAA 2799
Db |||||
QY 1618 GCTGAGTGTATCAAGGATCTCTGGCAGCCCTGGGCGATCTCTCATGACTCTCTCCAA 1677
Db |||||
QY 2800 GCTGAGTGTATCAAGGATCTCTGGCAGCCCTGGGCGATCTCTCATGACTCTCTCCAA 2859
Db |||||
QY 1678 GATCAGCTCGAGAAAGTCAAGGACTTCCAGAGAGAAATTCGCGCTCTGAAGAGAACGTG 1737
Db |||||
QY 2860 GATCAGCTCGAGAAAGTCAAGGACTTCCAGAGAGAAATTCGCGCTCTGAAGAGAACGTG 2919
Db |||||
QY 1738 AGCAGGTGATGACCTTGTCTGCCAGCTTACACCTTGGGCACTCAGCTCTCCAGTAT 1797
Db |||||
QY 2920 AGCAGGTGATGACCTTGTCTGCCAGCTTACACCTTGGGCACTCAGCTCTCCAGTAT 2979
Db |||||
QY 1798 AACCTCAGCACTCTGGAAGACCTTGAACCCAGATGGAAGCTTCTGAGGTGCCCTCGAG 1857
Db |||||
QY 2980 AACCTCAGCACTCTGGAAGACCTTGAACCCAGATGGAAGCTTCTGAGGTGCCCTCGAG 3039
Db |||||
QY 1858 GACGAGTCAAGGAGCTGATGAAGCCACAGGACTTTGGTCCAGCTCTCAGACCTT 1917
Db |||||
QY 3040 GACGAGTCAAGGAGCTGATGAAGCCACAGGACTTTGGTCCAGCTCTCAGACCTT 3099
Db |||||
QY 1918 CTTTCCAGCTCTGTCCAGGCTCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTAC 1977
Db |||||
QY 3100 CTTTCCAGCTCTGTCCAGGCTCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTAC 3159
Db |||||
QY 1978 TATATCAACCAAGAGACTCAAA 2001

Db 3160 TATATCAACCAAGAGACTCAAA 3183
|||
RESULT 9
AAD37255
ID AAD37255 standard; DNA; 5149 BP.
XX
AC AAD37255;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
XX WO200183695-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13677.
XX
XX 28-APR-2000; 2000US-200777P.
XX
XX (XIAO/) XIAO X.
XX
XX Xiao X;
XX
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -
XX
XX Example 1; Page 57-59; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid construct containing human dystrophin minigenes, a muscle
XX creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
XX Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 other;
XX
Query Match 83.1%; Score 1662.8; DB 24; Length 5149;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGAGCCGTTATCAACACGCTTTAGAGAA 60
Db 1757 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGAGCCGTTATCAACACGCTTTAGAGAA 1816
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAACAGAGAGATTTCTAAT 120
Db 1817 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTGCAAGCAACAGAGAGATTTCTAAT 1876
QY 121 GATGTGGAAGTGTGAAGACCAAGCTTTCTATCTCATGAGGGGTACATGATGGATTTGACA 180
Db 1877 GATGTGGAAGTGTGAAGACCAAGCTTTCTATCTCATGAGGGGTACATGATGGATTTGACA 1936

181 GCCCATCAGGCGCGGTGGTAAATTTCTACAAATGGGAAGTAAGCTGATGGACAGCA 240
1937 GCCCATCAGGCGCGGTGGTAAATTTCTACAAATGGGAAGTAAGCTGATGGACAGCA 1996
241 AAATTATCAGAAATGAGAAATCTGAAGTACAAGACAGATGAACTCTCTAATTCAGA 300
1997 AAATTATCAGAAATGAGAAATCTGAAGTACAAGACAGATGAACTCTCTAATTCAGA 2056
301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGATTTTAATG 360
2057 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGATTTTAATG 2116
361 GATCTCCAGAAATCAGAAATCTGAAGAGTGAATGACTGGCTTAACAAAAACAAGAAAGA 420
2117 GATCTCCAGAAATCAGAAATCTGAAGAGTGAATGACTGGCTTAACAAAAACAAGAAAGA 2176
421 ACAAGGAAATGAGGAAGAGCGCTCTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTA 480
2177 ACAAGGAAATGAGGAAGAGCGCTCTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTA 2236
481 CAACAACATTAAGTGTCTTCAAGAGATCTAGAAACAAGTCAAGGTCAATTCCTC 540
2237 CAACAACATTAAGTGTCTTCAAGAGATCTAGAAACAAGTCAAGGTCAATTCCTC 2296
541 ACTCAGATGCTGTGTAGTGTAGTGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAA 600
2297 ACTCAGATGCTGTGTAGTGTAGTGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAA 2356
601 GAACAACATTAAGTGTCTTCAAGAGATCTAGAAACAAGTCAAGGTCAATTCCTC 660
2357 GAACAACATTAAGTGTCTTCAAGAGATCTAGAAACAAGTCAAGGTCAATTCCTC 2416
661 TGGGTCTTTTACAGACAGCGCTGAC-----CTAGCTCCTGGACTGACCACT 708
2417 TGGGTCTTTTACAGACAGCGCTGAC-----CTAGCTCCTGGACTGACCACT 2476
709 ATTGGAGACCTCTCTACTCAGACTGTACTCTGGTGACACA-----ACCTGTGGTT 759
2477 TTTAGTGCATGGCTTTTCAAGAAAGAAGATGCAAGTGAACAAGATTCACACAAGCTGCTT 2536
760 ACTAAGAAATCGCATCTC-----779
2537 AAAGATCAAAATGAATGTTTATCAAGTCTTTCAAAAACCTGGCCGTTTAAAGCGGATCTA 2596
780 -----CAAACTAGAAATGCCATCTCC 801
2597 GAAAGAAAAGCAATCCATGGGCAAACTGTATCTACTCAACAAGATCTTCTTCAACA 2656
802 TTGATGTTGGAG-----813
2657 CTGAAGAATAAGTCAAGTGACCCAGAGACGGAAGCATGGCTGGATAACTTTGGCCCGGTGT 2716
814 -----GTACCTACTAGATTAAGTCA 837
2717 TGGGATAATTTAGTCCAAAACTTGAAGAAGATGACAGCAGACATGATTAAGTCA 2776
838 CAGTTCCTCCCTGGACCTGGAAAAAGTTCTTGGCTTGGCTTACAGAAGCTGAAAACTGCC 897
2777 CAGTTCCTCCCTGGACCTGGAAAAAGTTCTTGGCTTGGCTTACAGAAGCTGAAAACTGCC 2836
898 AATGCTCTCAGAGATGTACCCGTGAAGAAAGGCTCTAGAGACTTCAAGGAGTAAAA 957
2837 AATGCTCTCAGAGATGTACCCGTGAAGAAAGGCTCTAGAGACTTCAAGGAGTAAAA 2896
958 GAGCTGATGAACAATGGCAAGACCTCAAGGTGAATTAAGCTCACAGATGTTTAT 1017
2897 GAGCTGATGAACAATGGCAAGACCTCAAGGTGAATTAAGCTCACAGATGTTTAT 2956
1018 CACAACCTGGATGAACAAGCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGA 1077
2957 CACAACCTGGATGAACAAGCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGA 3016
1078 GTCCTGTTACAAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAACCTCGGAAAAAG 1137

3017 GTCTCTTACAAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAACCTCGAAAAAG 3076
1138 TCTCTCAACATTAAGTCCCATTTGGAAAGCCAGTCTTGACCAAGTGGAAAGCGTCTGCACCTT 1197
3077 TCTCTCAACATTAAGTCCCATTTGGAAAGCCAGTCTTGACCAAGTGGAAAGCGTCTGCACCTT 3136
1198 TCTCTGAGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGCA 1257
3137 TCTCTGAGGAACCTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGCAGCA 3196
1258 CCTATTGCGAGCGACTTTCCAGCAGATTCAGAAGCAGAACATGTACATAGGCGCTTCAAG 1317
3197 CCTATTGCGAGCGACTTTCCAGCAGATTCAGAAGCAGAACATGTACATAGGCGCTTCAAG 3256
1318 AGGAAATTGAATACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTT 1377
3257 AGGAAATTGAATACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTT 3316
1378 CTGACAGACGAGCGCTTTGGAGGACTAGAGAACTCTACCAAGAGCCCGCAGAGACTGCCT 1437
3317 CTGACAGACGAGCGCTTTGGAGGACTAGAGAACTCTACCAAGAGCCCGCAGAGACTGCCT 3376
1438 CCTGAGGAGAGAGCCCGCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAAT 1497
3377 CCTGAGGAGAGAGCCCGCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAAT 3436
1498 ACTGAGTGGGAAAAATTTGAACCTGCTGCTGACTGCGCAGAGAAAAATAGATGAGACC 1557
3437 ACTGAGTGGGAAAAATTTGAACCTGCTGCTGACTGCGCAGAGAAAAATAGATGAGACC 3496
1558 CTTGAAAGACTCTCAGGAACCTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCAA 1617
3497 CTTGAAAGACTCTCAGGAACCTTCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGCGCCAA 3556
1618 GCTGAGTGTATCAAGGATCTCTGCGAGCCCGTGGCGATCTCTCTATTGACTCTCTCAA 1677
3557 GCTGAGTGTATCAAGGATCTCTGCGAGCCCGTGGCGATCTCTCTATTGACTCTCTCAA 3616
1678 GATCACTCTGAGGAAGTCAAGGACTCTTGGAGGAAATTTGGGCTCTTGAAGAGAAAGTG 1737
3617 GATCACTCTGAGGAAGTCAAGGACTCTTGGAGGAAATTTGGGCTCTTGAAGAGAAAGTG 3676
1738 AGCCACGCTCAATGACCTTGTCTGCGAGCTTACCACTTTGGGCAATTCAGCTCTACCCGTAT 1797
3677 AGCCACGCTCAATGACCTTGTCTGCGAGCTTACCACTTTGGGCAATTCAGCTCTACCCGTAT 3736
1798 AACCTCAGACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGGTGGCGCTGAG 1857
3737 AACCTCAGACTCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGCAAGGTGGCGCTGAG 3796
1858 GACCGAGTCAGGAGCTGCAATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGACATTT 1917
3797 GACCGAGTCAGGAGCTGCAATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGACATTT 3856
1918 CTTTCCAGCTCTGTCAGGCTCCCTGGAGAGAGCCATCTGCGCCAAAACAAAGTCCCTTAC 1977
3857 CTTTCCAGCTCTGTCAGGCTCCCTGGAGAGAGCCATCTGCGCCAAAACAAAGTCCCTTAC 3916
1978 TATATCAACACGAGACTCAAAACA 2001
3917 TATATCAACACGAGACTCAAAACA 3940

RESULT 10

ABK81999

ID ABK81999 standard; DNA; 5462 BP.

XX ABK81999;

AC ABK81999;

XX 13-AUG-2002 (first entry)

XX 13-AUG-2002 (first entry)

XX 13-AUG-2002 (first entry)

DE DNA encoding mini-dystrophin protein deltaR2-R21+H3.

XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO200229056-A2.
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-US31126.
XX
XX 06-OCT-2000; 2000US-238848P.
XX
XX (UNMI) UNIV MICHIGAN.
XX
XX Chamberlain JS, Harper SQ;
XX
XX WPI; 2002-435334/46.
XX
XX A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
XX Disclosure; Fig 14; 145pp; English.
XX
XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX
XX Sequence 5462 BP; 1668 A; 1225 C; 1212 G; 1357 T; 0 other;
SQ
Query Match 64.1%; Score 1283; DB 24; Length 5462;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1302; Conservative 0; Mismatches 0; Indels 9; Gaps 1;
691 GCTCTGGACTGACCACTATTGGAGCCCTCTCTACTCTGACGCTTACTCTGGTGACAAA 750
1547 GCTCTGGACTGACCACTATTGGAGCCCTCTCTACTCTGACGCTTACTCTGGTGACAAA 1606
751 CTTGTGGTACTAAGGAACTGCCATCTCCAACTAGAAATGCCATCTTCTTGTATGTTG 810
1607 CTTGTGGTACTAAGGAACTGCCATCTCCAACTAGAAATGCCATCTTCTTGTATGTTG 1666
811 GAGGTACTCTACTCATAGATTACTGGAACAGTTCCCTCGACCTGGAAGTTCTTCTGCC 870
1667 GAG-----CATAGATTACTGCAACAGTTCCCTCGACCTGGAAGTTCTTCTGCC 1717
871 TGCGTTACAGAAGCTGAAACAACTGCCAATGTCTACAGAGTGTACCGTGAAGAAAGG 930
1718 TGCGTTACAGAAGCTGAAACAACTGCCAATGTCTACAGAGTGTACCGTGAAGAAAGG 1777
931 CTCTAGAGACTCCAAGGAGTAAAGAGCTGATGAAACATGGCAAGCTCCAAGT 990
1778 CTCTAGAGACTCCAAGGAGTAAAGAGCTGATGAAACATGGCAAGCTCCAAGT 1837
991 GAAATTGAAGCTCACAGAGTGTATATCAACCTGGATGAAACAGCCAAAATTCCTG 1050
1838 GAAATTGAAGCTCACAGAGTGTATATCAACCTGGATGAAACAGCCAAAATTCCTG 1897
1051 AGATCCCTGGAAGTTCCTGATGATGAGTCTCTGTTACAAAGAGCTTTGGATTAACATGAAC 1110
1998 AGATCCCTGGAAGTTCCTGATGATGAGTCTCTGTTACAAAGAGCTTTGGATTAACATGAAC 1957
1111 TTCAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGGAAGCCAGT 1170

Db 1958 TTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCATTTTGGAAAGCCAGT 2017
QY 1171 TCTGACCACTGGAGCGTCTGACCTTCTCTGAGGAACCTTCTGGTGTGCTGACAGCTG 1230
Db 2018 TCTGACCACTGGAGCGTCTGACCTTCTCTGAGGAACCTTCTGGTGTGCTGACAGCTG 2077
QY 1231 AAAAGATGATGAATTAAAGCCGCGAGCCACTATTGGAGCGGACCTTCCAGAGCTTCAAGAG 1290
Db 2078 AAAAGATGATGAATTAAAGCCGCGAGCCACTATTGGAGCGGACCTTCCAGAGCTTCAAGAG 2137
QY 1291 CAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACTTAAAGAACCTGTATCATG 1350
Db 2138 CAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACTTAAAGAACCTGTATCATG 2197
QY 1351 AGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCGCTTTTGGAAAGGACTAGAGAAA 1410
Db 2198 AGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCGCTTTTGGAAAGGACTAGAGAAA 2257
QY 1411 CTCTACAGAGCGCCAGAGAGCTGCTCTCTGAGGAGAGCGCCAGAAATGTCTACCTCGGCTT 1470
Db 2258 CTCTACAGAGCGCCAGAGAGCTGCTCTCTGAGGAGAGCGCCAGAAATGTCTACCTCGGCTT 2317
QY 1471 CTACGAAAGCAGGCTGAGAGGCTCAATATCTGAGTGGGAAAAATTTGAACTGCACTCCGCT 1530
Db 2318 CTACGAAAGCAGGCTGAGAGGCTCAATATCTGAGTGGGAAAAATTTGAACTGCACTCCGCT 2377
QY 1531 GACTGCGAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCAG 1590
Db 2378 GACTGCGAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCAG 2437
QY 1591 GATGAGCTGGACTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCTG 1650
Db 2438 GATGAGCTGGACTCAAGCTGGCCAAAGCTGAGGTGATCAAGGGATCTTGGCAGCCCTG 2497
QY 1651 GCGATCTCTCTATTGACTCTCTCAAGATCACTCTGAGAAAGTCAAGGCACTTTCGAGA 1710
Db 2498 GCGATCTCTCTATTGACTCTCTCAAGATCACTCTGAGAAAGTCAAGGCACTTTCGAGA 2557
QY 1711 GAAATTCGCGCTCTGAAAGAGAACTGAGCCACCTCAATGACCTTGTCTGCGCAGCTTACC 1770
Db 2558 GAAATTCGCGCTCTGAAAGAGAACTGAGCCACCTCAATGACCTTGTCTGCGCAGCTTACC 2617
QY 1771 ACTTTGGCATTGAGCTCTCAAGCTGATTAACCTCAGCACTCTCTGGAAGACTTGAACCCAGA 1830
Db 2618 ACTTTGGCATTGAGCTCTCAAGCTGATTAACCTCAGCACTCTCTGGAAGACTTGAACCCAGA 2677
QY 1831 TGAAGCTTCTGAGCTGGCCGCTGAGGACCGAGTCAGGACCTGATGAAGCCACAG 1890
Db 2678 TGAAGCTTCTGAGCTGGCCGCTGAGGACCGAGTCAGGACCTGATGAAGCCACAG 2737
QY 1891 GACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTGTCAGGCTCCCTGGGAGAGA 1950
Db 2738 GACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTGTCAGGCTCCCTGGGAGAGA 2797
QY 1951 GCGATCTGCGCAACAAAGTGGCCCTACTATATCAACCCAGGACTCAAAACA 2001
Db 2798 GCGATCTGCGCAACAAAGTGGCCCTACTATATCAACCCAGGACTCAAAACA 2848
RESULT 11
AAD06794
ID AAD06794 standard; DNA; 5952 BP.
XX
AC AAD06794;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human dystrophin gene (Becker form).
XX
KW Human; dystrophin; extein; intein; trans-splicing; gene therapy;
KW Duchenne muscular dystrophy; Becker muscular dystrophy; DMD; BMD; ds.
XX

[illegible]

QY	993	AAATTGAAGCTCACACAGATGTTTATCACACCTCGGATGAAACAGCCAAATAATCTGAG	1053
DB	3172	AAATTGAAGCTCACACAGATGTTTATCACACCTCGGATGAAACAGCCAAATAATCTGAG	3233
QY	1053	ATCCTCGGAAGGTTCCGATGATCGAGTCTCTGTTTACAAAGACGTTTCTGATTAACATGAAC	1112
DB	3232	ATCCTCGGAAGGTTCCGATGATCGAGTCTCTGTTTACAAAGACGTTTCTGATTAACATGAAC	3291
QY	1113	CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAAGTCTCCATTTTGAAGCCAGTTC	1172
DB	3292	CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAAGTCTCCATTTTGAAGCCAGTTC	3351
QY	1173	TGACCAAGTGAAGGCTTCGACCTTTCTCTGAGGAGACTTCTGTTGTTGCTCAGCTGAA	1232
DB	3352	TGACCAAGTGAAGGCTTCGACCTTTCTCTGAGGAGACTTCTGTTGTTGCTCAGCTGAA	3411
QY	1233	AGATCATGAATTAAGCCCGGAGGACCTATTGGAGGCGACTTTTCCAGCAGTTTCCAGAGCA	1292
DB	3412	AGATCATGAATTAAGCCCGGAGGACCTATTGGAGGCGACTTTTCCAGCAGTTTCCAGAGCA	3471
QY	1293	GAAAGATGATACATAGGCGCTTCAAGAGGGAATTCAAAACCTTAAAGAACCTCTATCATGAG	1352
DB	3472	GAAAGATGATACATAGGCGCTTCAAGAGGGAATTCAAAACCTTAAAGAACCTCTATCATGAG	3531
QY	1353	TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTTGAAGGACTTAGAGAACT	1412
DB	3532	TACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCCTTTTGAAGGACTTAGAGAACT	3591
QY	1413	CTACAGAGGCGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCACAGATCTCACTCGGCTTCT	1472
DB	3592	CTACAGAGGCGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCACAGATCTCACTCGGCTTCT	3651
QY	1473	ACGAAAGCAGGCTGAGGAGGTCATATCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA	1532
DB	3652	ACGAAAGCAGGCTGAGGAGGTCATATCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGA	3711
QY	1533	CTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACTTCAAGAGGCGCACGGA	1592
DB	3712	CTGGCAGAGAAAAATAGATGAGACCTTGAAGACTTCAAGGAACTTCAAGAGGCGCACGGA	3771
QY	1593	TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGCATCAAGGATCTTGGGAGCCCGTGGG	1652
DB	3772	TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGCATCAAGGATCTTGGGAGCCCGTGGG	3831
QY	1653	CGATCTCTCATTTGACTCTCTCCAGATCACCTCGAGAACTCAAGGCACCTTCGAGGAGA	1712
DB	3832	CGATCTCTCATTTGACTCTCTCCAGATCACCTCGAGAACTCAAGGCACCTTCGAGGAGA	3891
QY	1713	AAATTGGGCTCTGAAAGAGAACGTGAGCCAGTCAATGACCTTCTCGCCAGCTTACAC	1772
DB	3892	AAATTGGGCTCTGAAAGAGAACGTGAGCCAGTCAATGACCTTCTCGCCAGCTTACAC	3951
QY	1773	TTTGGGCATTGAGCTCTCACCGTATTAACCTCAGCAGCTCTGGAAGACCTTGAACACAGATG	1832
DB	3952	TTTGGGCATTGAGCTCTCACCGTATTAACCTCAGCAGCTCTGGAAGACCTTGAACACAGATG	4011
QY	1833	GAAGCTTTCTGCAAGTGGGCGTTCGAGGACCGAGTCAAGGACTCAAGGCACCTTCAGGAGA	1892
DB	4012	GAAGCTTTCTGCAAGTGGGCGTTCGAGGACCGAGTCAAGGACTCAAGGCACCTTCAGGAGA	4071
QY	1893	CTTTGGTTCAGGATCTCAGCAGCTTTCTTTTCCACGCTCTGTCAGAGGTCCTCTGGGAGAGC	1955
DB	4072	CTTTGGTTCAGGATCTCAGCAGCTTTCTTTTCCACGCTCTGTCAGAGGTCCTCTGGGAGAGC	4133
QY	1953	CATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAAAACA	2001
DB	4132	CATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAAAACA	4180
RESULT 13			
AAD37229			
ID	AAD37229 standard; DNA; 11058 BP.		
XX			

RESULT 13
AAD37229
ID AAD37229 standard; DNA; 11058 BP.
XX

1	CC	AAD37229;	Db	8352	CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCTTGGAAAGCCAGTTC	8411
2	CC	21-AUG-2002 (first entry)	Qy	1173	TGACCAAGTGGAAAGCGTCTGCACCTTCTCTGAGGAACTTCTGGTGTGCTACAGCTGAA	1232
3	CC	Human dystrophin protein coding sequence.	Db	8412	TGACCAAGTGGAAAGCGTCTGCACCTTCTCTGAGGAACTTCTGGTGTGCTACAGCTGAA	8471
4	CC	Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;	Qy	1233	AGATGATGAATTAAGCCGGCAGCAGCTTATTGGAGCGAGCTTCCAGCAGTTCAGAGCA	1292
5	CC	adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;	Db	8472	AGATGATGAATTAAGCCGGCAGCAGCTTATTGGAGCGAGCTTCCAGCAGTTCAGAGCA	8531
6	CC	Becker muscular dystrophy; ds.	Qy	1293	GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG	1352
7	CC	Homo sapiens.	Db	8532	GAACGATGTACATAGGCGCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTATCATGAG	8591
8	CC	WO200183695-A2.	Qy	1353	TACTCTTGAGACTGTACGAATATTCTGACAGAGCCCTTTGGAAGGACTAGAGAACT	1412
9	CC	08-NOV-2001.	Db	8592	TACTCTTGAGACTGTACGAATATTCTGACAGAGCCCTTTGGAAGGACTAGAGAACT	8651
10	CC	27-APR-2001; 2001WO-US13677.	Qy	1413	CTACCAAGGAGCCAGAGAGCTGCTCTGAGAGAGAGAGCCAGAAATGTCACTCGGCTTCT	1472
11	CC	28-APR-2000; 2000US-200777P.	Db	8652	CTACCAAGGAGCCAGAGAGCTGCTCTGAGAGAGAGAGCCAGAAATGTCACTCGGCTTCT	8711
12	CC	(XIAO/) XIAO X.	Qy	1473	ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCCTCCCTGA	1532
13	CC	Xiao X;	Db	8712	ACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCCTCCCTGA	8771
14	CC	WPI; 2002-049342/06.	Qy	1533	CTGCGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGA	1592
15	CC	New dystrophin minigene for treating Duchenne or Becker muscular	Db	8772	CTGCGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAACTTCAAGAGGCCACGGA	8831
16	CC	dystrophy comprises an N-terminal domain or modified N-terminal domain,	Qy	1593	TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTGGCAGCCCTGGG	1652
17	CC	rod repeats, H1 and H4 domains and a cysteine rich domain of a	Db	8832	TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTGGCAGCCCTGGG	8891
18	CC	dystrophin gene -	Qy	1653	CGATCTCTCATTTGACTCTCTCCAGATCACTCCGAGAAAGTCAAGGCACTTCGAGGAGA	1712
19	CC	Example 1; Page 40-43; 71pp; English.	Db	8892	CGATCTCTCATTTGACTCTCTCCAGATCACTCCGAGAAAGTCAAGGCACTTCGAGGAGA	8951
20	CC	The present invention relates to an isolated nucleotide sequence encoding	Qy	1713	AATTGCGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTCGCCAGCTTACCAC	1772
21	CC	a dystrophin minigene. The minigene comprises N-terminal or modified	Db	8952	AATTGCGCCCTCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTCGCCAGCTTACCAC	9011
22	CC	N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4	Qy	1773	TTTGGGCAATTCAGCTCTCAGCGTATTAACCTCAGACTCTGGAAGACTGGAACACCATG	1832
23	CC	domains and cysteine-rich domains of dystrophin or utrophin genes. The	Db	9012	TTTGGGCAATTCAGCTCTCAGCGTATTAACCTCAGACTCTGGAAGACTGGAACACCATG	9071
24	CC	invention also relates to a recombinant adeno-associated virus (AAV)	Qy	1833	GAAGCTTCTGAGGTGCGCCGTCGAGGACCGAGTCAGGACCTGCATGAAGCCACAGGGA	1892
25	CC	comprising dystrophin minigene operably linked to an expression control	Db	9072	GAAGCTTCTGAGGTGCGCCGTCGAGGACCGAGTCAGGACCTGCATGAAGCCACAGGGA	9131
26	CC	element. The dystrophin minigene in operable linkage with an expression	Qy	1893	CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCCAGGTCCTTGGAGAGAGC	1952
27	CC	control element, in a recombinant adeno-associated virus or retrovirus is	Db	9132	CTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTGTGTCCAGGTCCTTGGAGAGAGC	9191
28	CC	useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular	Qy	1953	CATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAAAACA	2001
29	CC	dystrophy (BMD) in a mammalian subject. The present sequence is human	Db	9192	CATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAAAACA	9240
30	CC	dystrophin protein coding sequence.				
31	CC	Sequence 11058 BP; 3686 A; 2292 C; 2621 G; 2459 T; 0 other;				
32	CC	Query Match 59.1%; Score 1182.6; DB 24; Length 11058;				
33	CC	Best Local Similarity 99.7%; Pred. No. 0;				
34	CC	Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;				
35	CC	813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTTGGACCTGGAAAGTTCTTGGCTG 872	Qy			
36	CC	8052 GGAAGAAACTCATAGATTACTGCAACAGTTCCCTTGGACCTGGAAAGTTCTTGGCTG 8111	Db			
37	CC	873 GCTTACAGAGCTGAAACAACTGCAATGTCCTACAGGATGCTACCGTAAGGAAAGGCT 932	Qy			
38	CC	8112 GCTTACAGAGCTGAAACAACTGCAATGTCCTACAGGATGCTACCGTAAGGAAAGGCT 8171	Db			
39	CC	933 CCTAGAGACTTCCAGGAGTAAAGAGCTGATGAACAAATGGCAGAGCTTCCAGGTGA 992	Qy			
40	CC	8172 CCTAGAGACTTCCAGGAGTAAAGAGCTGATGAACAAATGGCAGAGCTTCCAGGTGA 8231	Db			
41	CC	993 AATTGAGCTCACACAGATGTTTATCAACCTGGATGAAACAGCCAAAATCTCTGAG 1052	Qy			
42	CC	8232 AATTGAGCTCACACAGATGTTTATCAACCTGGATGAAACAGCCAAAATCTCTGAG 8291	Db			
43	CC	1053 ATCCCTGGAAGTTCCGATGATGAGTCTGTTTCAAGAGCTTTGATGAATGAATCACTT 1112	Qy			
44	CC	8292 ATCCCTGGAAGTTCCGATGATGAGTCTGTTTCAAGAGCTTTGATGAATGAATCACTT 8351	Db			
45	CC	1113 CAGTGGAGTGAACCTCGAAAAAGTCTTCAACATTAGGTCCCATTTGGAAGCAGTTC 1172	Qy			

RESULT 14
ABK82005
ID ABK82005 standard; DNA; 11241 BP.
XX
AC ABK82005;
XX
DT 13-AUG-2002 (first entry)
XX
DE cDNA encoding human dystrophin, full length HDMD.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; human; ds.
XX
OS Homo sapiens.

```
XX PN WO200229056-A2.
XX PD 11-APR-2002.
XX PF 04-OCT-2001; 2001WO-US311126.
XX PR 06-OCT-2000; 2000US-238848P.
XX PA (UNMI ) UNIV MICHIGAN.
XX PI Chamberlain JS, Harper SQ;
XX DR WPI; 2002-435334/46.
XX PT A composition for preparing therapeutic drugs, has a mini-dystrophin
XX PT peptide comprising a specific number of spectrin-like repeat domains,
XX PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX PS Example 2; Fig 23; 145pp; English.
XX CC The invention describes a composition comprising a mini-dystrophin
XX CC peptide comprising a spectrin-like repeat domain, where the domain
XX CC comprises n spectrin-like repeats, and contains no more than n
XX CC spectrin-like repeats, where n is an even number between 4-24, or a
XX CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
XX CC peptide or the polynucleotide encoding it is useful as a medicament,
XX CC for preparing a drug for therapeutic application and in the preparation
XX CC of a composition for treatment of muscle disease, e.g. Duchenne's
XX CC muscular dystrophy (DMD). This sequence represents a human dystrophin
XX CC polynucleotide sequence used in the creation of the mini-dystrophin
XX CC peptides of the invention.
XX SQ Sequence 11241 BP; 3738 A; 2325 C; 2656 G; 2522 T; 0 other;

Query Match 59.1%; Score 1182.6; DB 24; Length 11241;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTCAACAGTTCCTCCCTGACCTGGAAAGATTTCTTGCTG 872
DB 8260 GGAAGAACTCATAGATTACTCAACAGTTCCTCCCTGACCTGGAAAGATTTCTTGCTG 8319
QY 873 GCTTACAGAGCTGAACACTGCCAATCTCTTACAGGATGCTACCTCCCTTAAGGAAGGCT 932
DB 8320 GCTTACAGAGCTGAACACTGCCAATCTCTTACAGGATGCTACCTCCCTTAAGGAAGGCT 8379
QY 933 CTTAGAGACTCCAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAAGTGA 992
DB 8380 CTTAGAGACTCCAGGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAAGTGA 8439
QY 993 AATTGAAGCTCACAGATGTTTATCACAACTGATGAACACAGCCAAACAAATCCCTGAG 1052
DB 8440 AATTGAAGCTCACAGATGTTTATCACAACTGATGAACACAGCCAAACAAATCCCTGAG 8499
QY 1053 ATCCCTGGAAGGTTCCGATGATGCTGCTGTTACAAAGACGTTTGGATTAACATGAAT 1112
DB 8500 ATCCCTGGAAGGTTCCGATGATGCTGCTGTTACAAAGACGTTTGGATTAACATGAAT 8559
QY 1113 CAAGTCGAGTGAATTTGGAAAGGCTCTCTCAACATTAGGTCCCATTTGGAAAGCAGTTC 1172
DB 8560 CAAGTCGAGTGAATTTGGAAAGGCTCTCTCAACATTAGGTCCCATTTGGAAAGCAGTTC 8619
QY 1173 TGACAGTGGAGCGTCTGACCTTTCTCTGAGGAGAACTTCTGTTGGCTACAGCTGAA 1232
DB 8620 TGACAGTGGAGCGTCTGACCTTTCTCTGAGGAGAACTTCTGTTGGCTACAGCTGAA 8679
QY 1233 AGATGATGAATTAAGCCGAGGACCTATTGGAGCGACTTTCCAGCAGTTCAGAGCA 1292
DB 8680 AGATGATGAATTAAGCCGAGGACCTATTGGAGCGACTTTCCAGCAGTTCAGAGCA 8739
QY 1293 GAACGATGATACATAGGCTTCAAGAGGGAATTTGAACATTAAGACCTGTAATCATGAG 1352
DB 8740 GAACGATGATACATAGGCTTCAAGAGGGAATTTGAACATTAAGACCTGTAATCATGAG 8799
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTTGAAGGACTAGAGAACT 1412
DB 8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTTGAAGGACTAGAGAACT 8859
QY 1413 CTACCAGAGCCCGAGAGAGCTGCTCTCTGAGAGAGAGCCCGAGATGTCATCTCGGCTTCT 1472
DB 8860 CTACCAGAGCCCGAGAGAGCTGCTCTCTGAGAGAGAGCCCGAGATGTCATCTCGGCTTCT 8919
QY 1473 ACGAAAGCAGGCTGAGGAGGTCATATCTGAGTGGGAAAAATTTGAACCTGCACTCGCTGA 1532
DB 8920 ACGAAAGCAGGCTGAGGAGGTCATATCTGAGTGGGAAAAATTTGAACCTGCACTCGCTGA 8979
QY 1533 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAATTTCAAGAGGCCACCGA 1592
DB 8980 CTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGGAATTTCAAGAGGCCACCGA 9039
QY 1593 TGAGCTGGACCTCAAGCTGGCCAGCTGAGTGTATCAAGGATCTCTGGCAGCCGCTGGG 1652
DB 9040 TGAGCTGGACCTCAAGCTGGCCAGCTGAGTGTATCAAGGATCTCTGGCAGCCGCTGGG 9099
QY 1653 CGATCTCTCTCATTTGACTCTCTCTCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
DB 9100 CGATCTCTCTCATTTGACTCTCTCTCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGA 9159
QY 1713 AATTGGGCTCTGAAGAGAGAGTCAAGGAGTCAATGACCTTGTCTCCAGCTTACCAC 1772
DB 9160 AATTGGGCTCTGAAGAGAGAGTCAAGGAGTCAATGACCTTGTCTCCAGCTTACCAC 9219
QY 1773 TTTGGCATTTCAGCTCTCACCGTATTAACCTTCAGCACTCTGGAAGACCTGAACACCATG 1832
DB 9220 TTTGGCATTTCAGCTCTCACCGTATTAACCTTCAGCACTCTGGAAGACCTGAACACCATG 9279
QY 1833 GAAGCTTTCGAGGCTGCGCTCGAGGACCGAGTCAAGGAGTCAAGGCTGATGAAGCCACAGGGA 1892
DB 9280 GAAGCTTTCGAGGCTGCGCTCGAGGACCGAGTCAAGGAGTCAAGGCTGATGAAGCCACAGGGA 9339
QY 1893 CTTTGTCTCAGCATCTCAGCACTTTTCTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGC 1952
DB 9340 CTTTGTCTCAGCATCTCAGCACTTTTCTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGC 9399
QY 1953 CATCTCGCAAAACAAAGTGCCTTACTATATCAACACAGAGACTCAACA 2001
DB 9400 CATCTCGCAAAACAAAGTGCCTTACTATATCAACACAGAGACTCAACA 9448

RESULT 15
ABK82002
ID ABK82002 standard; DNA; 11443 BP.
XX AC ABK82002;
XX XX
XX DT 13-AUG-2002 (first entry)
XX DE DNA encoding mini-dystrophin protein deltaR9-R16.
XX XX
XX KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
XX KM Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX XX
XX PN WO200229056-A2.
XX PD 11-APR-2002.
XX PF 04-OCT-2001; 2001WO-US311126.
XX PR 06-OCT-2000; 2000US-238848P.
XX PA (UNMI ) UNIV MICHIGAN.
```


PI Chamberlain JS, Harper SQ;

XX WPI; 2002-435334/46.

XX A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -

XX Disclosure; Fig 17; 145pp; English.

XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.

XX Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 other;

Query Match 59.1%; Score 1182.6; DB 24; Length 11443;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGTTCCTTGCCCTG 872

DB 5746 GGAAGAACTACTAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGTTCCTTGCCCTG 5805

QY 873 GCTTACAGAAAGCTGAACAACTGCCAATGCTCTACAGGATGCTACCGTAAGAAAGGCT 932

DB 5806 GCTTACAGAAAGCTGAACAACTGCCAATGCTCTACAGGATGCTACCGTAAGAAAGGCT 5865

QY 933 CCTAGAGACTCCCAAGGAGTAAAGAGCTGTGTAACAACTGGCAAGACTCCCAAGGTGA 992

DB 5866 CCTAGAGACTCCCAAGGAGTAAAGAGCTGTGTAACAACTGGCAAGACTCCCAAGGTGA 5925

QY 993 AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAGCCAAATAATCCTGAG 1052

DB 5926 AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAGCCAAATAATCCTGAG 5985

QY 1053 ATCCCTGGAGGTTCCGATGCTGAGTCCCTGTTACAAAGAGCTTTCGTAAGATGAATTT 1112

DB 5986 ATCCCTGGAGGTTCCGATGCTGAGTCCCTGTTACAAAGAGCTTTCGTAAGATGAATTT 6045

QY 1113 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCTATTGGAAGCCAGTTC 1172

DB 6046 CAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCTATTGGAAGCCAGTTC 6105

QY 1173 TGACCACTGGAGGCTGTCACCTTTCTCTGAGGAACTTCTGTTGGCTACAGCTGAA 1232

DB 6106 TGACCACTGGAGGCTGTCACCTTTCTCTGAGGAACTTCTGTTGGCTACAGCTGAA 6165

QY 1233 AGATGATGAATTAAGCCGCGGAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCA 1292

DB 6166 AGATGATGAATTAAGCCGCGGAGGACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCA 6225

QY 1293 GAACGATGTACATAGGCTTTCAAGAGGAAATGAAACCTTAAGAACTGTATCATGAG 1352

DB 6226 GAACGATGTACATAGGCTTTCAAGAGGAAATGAAACCTTAAGAACTGTATCATGAG 6285

QY 1353 TACTCTTTGAGACTGTAGAAATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 1412

DB 6286 TACTCTTTGAGACTGTAGAAATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAACT 6345

QY 1413 CTACCAGAGCCCAAGAGCTGCTCTCTGAGGAGAGCCAGAACTGTCTCGCTTCT 1472

DB 6346 CTACCAGAGCCCAAGAGCTGCTCTCTGAGGAGAGCCAGAACTGTCTCGCTTCT 6405

QY 1473 ACGAAAGCAGCTGAGGAGTCAATCTGAGTGGGAATTTGAACTGCACTCCGCTGA 1532

DB 6406 ACGAAAGCAGCTGAGGAGTCAATCTGAGTGGGAATAATTGAACCTGCATCCGCTGA 6465

QY 1533 CTGGCAGAGAAATAGATGAGACCTTGAAGACTCCAGGAACCTTCAAGAGCCACGGA 1592

DB 6466 CTGGCAGAGAAATAGATGAGACCTTGAAGACTCCAGGAACCTTCAAGAGCCACGGA 6525

QY 1593 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTCTGGAGCCCGTGGG 1652

DB 6526 TGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTCTGGAGCCCGTGGG 6585

QY 1653 CGATCTCTCTATTGACTCTCTCCCAAGATCACTCGAGAAAGTCAAGSCACTTCGAGGAGA 1712

DB 6586 CGATCTCTCTATTGACTCTCTCCCAAGATCACTCGAGAAAGTCAAGSCACTTCGAGGAGA 6645

QY 1713 AATTGCGCCTCTGAAAGAGAAACCTGAGCCACGTCATGACCTTTGCTGCGCAGCTTACAC 1772

DB 6646 AATTGCGCCTCTGAAAGAGAAACCTGAGCCACGTCATGACCTTTGCTGCGCAGCTTACAC 6705

QY 1773 TTTGGGCAATTGACTCTCAAGCTGATTAACCTCAGCAGCTCTGGAAGACCTGAACACCATG 1832

DB 6706 TTTGGGCAATTGACTCTCAAGCTGATTAACCTCAGCAGCTCTGGAAGACCTGAACACCATG 6765

QY 1833 GAAGCTTTCTGCGAGTGGCCCTGCGAGGACCGAGTCAGGACCTGCATGAAGCCCAACAGGA 1892

DB 6766 GAAGCTTTCTGCGAGTGGCCCTGCGAGGACCGAGTCAGGACCTGCATGAAGCCCAACAGGA 6825

QY 1893 CTTTGGTCCAGCATCTCAGCAGCTTTCTTTCCAGCTCTGTCAGGCTCCCTGGAGAGAGC 1952

DB 6826 CTTTGGTCCAGCATCTCAGCAGCTTTCTTTCCAGCTCTGTCAGGCTCCCTGGAGAGAGC 6885

QY 1953 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACA 2001

DB 6886 CATCTCGCCAAACAAAGTGCCCTACTATATCAACCCAGAGACTCAAAACA 6934

Search completed: February 1, 2004, 12:04:09

Job time : 487.292 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

M nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:14:55 ; Search time 7128.21 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-6_COPY_1000_3000
Perfect score: 2001
Sequence: 1 ggcagttcatgatgagag.....tcaaccacgagactcaaca 2001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2889711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_hgtg_hum.*
- 40: em_hgtg_mus.*
- 41: em_hgtg_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1283	64.1	5462	6	AX538621 Sequence
2	1182.6	59.1	5952	6	AR304538 Sequence
3	1182.6	59.1	5952	6	AX114289 Sequence
4	1182.6	59.1	8659	6	AX538622 Sequence
5	1182.6	59.1	11443	6	AX538624 Sequence
6	1182.6	59.1	12057	6	AX538627 Sequence
7	1182.6	59.1	12446	9	HSDMDR
8	1182.6	59.1	13957	6	AX409637 Sequence
9	1182.6	59.1	13957	6	AX538581 Sequence
10	1182.6	59.1	13957	9	HUMDYS
11	1182.6	59.1	13977	6	AR220819 Sequence
12	1180	59.0	5339	6	AX538620 Sequence
13	1030.6	51.5	13887	4	AF070485
14	1004	50.2	13815	6	AX306153 Sequence
15	1004	50.2	13815	6	AX538582 Sequence
16	1004	50.2	13815	10	MUSDYSA
17	1004	50.2	19307	6	AR033392 Sequence
18	1004	50.2	19307	6	AR142592 Sequence
19	765	38.2	13575	5	GGDYS
20	727	36.3	5417	6	AX538619 Sequence
21	678.4	33.9	4402	6	E30220
22	618.8	30.9	4075	6	E30221
23	614.2	30.7	4402	6	E30219
24	547.4	27.4	3275	10	MUSDYS
25	459.4	23.0	650	9	HSDMDP1
26	422.6	21.1	3521	5	AF339031
27	408.2	20.4	11096	6	AX538584 Sequence
28	408.2	20.4	11096	10	MMY12229
29	408	20.4	10705	10	RNAJ2967
30	397	19.8	6045	6	A63605 Sequence
31	397	19.8	6045	6	AR281528 Sequence
32	397	19.8	6059	6	AX107972 Sequence
33	397	19.8	10302	6	AX538583 Sequence
34	397	19.8	10302	9	HSMUFS
35	397	19.8	10320	6	A63607
36	397	19.8	10320	6	AR281529 Sequence
37	391.8	19.6	2654	5	FSCDYSTRO
38	387	19.3	387	6	AX538612 Sequence
39	378.6	18.9	3161	10	MMGUTRPH
40	348	17.4	348	6	AX538611 Sequence
41	330	16.5	333	6	AX538589 Sequence
42	327	16.3	327	6	AX538588 Sequence
43	324	16.2	324	6	AX538613 Sequence
44	322.6	16.1	3499	9	HSU43519
45	322.6	16.1	5106	6	AX552248 Sequence

ALIGNMENTS

RESULT 1	AX538621	Sequence 41 from Patent WO0229056.	5462 bp	DNA	linear	PAT 23-NOV-2002
AX538621	LOCUS	AX538621				
DEFINITION	Sequence 41 from Patent WO0229056.					
ACCESSION	AX538621					
VERSION	AX538621.1	GI:25271168				
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	artificial sequences.					
REFERENCE	1					
AUTHORS	Chamberlain, J.S. and Harper, S.Q.					
TITLE	Mini-dystrophin nucleic acid and peptide sequences					
JOURNAL	Patent: WO 0229056-A 41 11-APR-2002;					
	THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)					

[illegible]

1113 CAAGTGGAGTGAATTCGGAAAGTCTCTACATTTAGTCCCATTTGGAAAGCCAGTTC 1172
3246 CAAGTGGAGTGAATTCGGAAAGTCTCTACATTTAGTCCCATTTGGAAAGCCAGTTC 3305
1173 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAAGGAATCTTGTGTGGCTACAGTGA 1232
3306 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGCAAGGAATCTTGTGTGGCTACAGTGA 3365
1233 AGATGATGAATTAAGCGGCGAGCCTATTGAGGCGGACTTTCAGAGCAATTCAGAGCA 1292
3366 AGATGATGAATTAAGCGGCGAGCCTATTGAGGCGGACTTTCAGAGCAATTCAGAGCA 3425
1293 GAACGATGATACATAGGCGCTTCAAGAGGAATTTGAAAACTTAAGAACCTGTATCATGAG 1352
3426 GAACGATGATACATAGGCGCTTCAAGAGGAATTTGAAAACTTAAGAACCTGTATCATGAG 3485
1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGAGCCCTTTGGAAGGACTAGAGAACT 1412
3486 TACTCTTGAGACTGTACGAATATTTCTGACAGAGAGCCCTTTGGAAGGACTAGAGAACT 3545
1413 CTACAGAGCCAGAGAGTGTCTCTGAGGAGAGAGCCAGAAATGTCTACCTGGCTTCT 1472
3546 CTACAGAGCCAGAGAGTGTCTCTGAGGAGAGAGCCAGAAATGTCTACCTGGCTTCT 3605
1473 ACGAAAGCGCTGAGAGGTCATATCTGAGTGGGAAAAATTTGAACCTGCTCCGCTGA 1532
3606 ACGAAAGCGCTGAGAGGTCATATCTGAGTGGGAAAAATTTGAACCTGCTCCGCTGA 3665
1533 CTGACAGAGAAATATGATGAGAGCCCTTTGAAAGACTCCAGAGAACTTCAAGAGCCACGGA 1592
3666 CTGACAGAGAAATATGATGAGAGCCCTTTGAAAGACTCCAGAGAACTTCAAGAGCCACGGA 3725
1593 TGAGCTGAGCCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGAGATCTTGGCAGCCCTGGG 1652
3726 TGAGCTGAGCCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGAGATCTTGGCAGCCCTGGG 3785
1653 CGATCTCTCATTTGACTCTCTCAAGATCACTTCGAGAAAGTCAAGGCACTTCGAGGAGA 1712
3786 CGATCTCTCATTTGACTCTCTCAAGATCACTTCGAGAAAGTCAAGGCACTTCGAGGAGA 3845
1713 AATTGGCGCTCTGAAAGAGAACGTGAGCCAGTCAATGACCTTGTCTCGCAGCTTACCAC 1772
3846 AATTGGCGCTCTGAAAGAGAACGTGAGCCAGTCAATGACCTTGTCTCGCAGCTTACCAC 3905
1773 TTTGGGCAATTCAGTCTCACCCTATACCTCAGCACTCTGAGAGACCTGAACCCAGATG 1832
3906 TTTGGGCAATTCAGTCTCACCCTATACCTCAGCACTCTGAGAGACCTGAACCCAGATG 3965
1833 GAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGGAGTCAATGAGCTGATCAAGCCAGGGA 1892
3966 GAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGGAGTCAATGAGCTGATCAAGCCAGGGA 4025
1893 CTTTGGTCCAGCATCTCAGCACTTTCTTTTCAAGTCTGAGGAGTCCCTGGGAGAGGC 1952
4026 CTTTGGTCCAGCATCTCAGCACTTTCTTTTCAAGTCTGAGGAGTCCCTGGGAGAGGC 4085
1953 CATCTCGCCAAACAAAGTCCCTACTATATCAACACAGAGACTCAAAACA 2001
4086 CATCTCGCCAAACAAAGTCCCTACTATATCAACACAGAGACTCAAAACA 4134

RESULT 3
AX114289
LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent WO0129243.
ACCESSION AX114289
VERSION AX114289.1 GI:14031259
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Paul,X.L. and Xiao,X.
TITLE Method and vector for producing and transferring trans -spliced peptides
JOURNAL Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
SYSTEM OF HIGHER EDUCATION (US)
FEATURES
Location/Qualifiers
source 1..5952
misc_feature 2897..2898
misc_feature 3198..3199
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 59.1%; Score 1182.6; DB 6; Length 5952;
Best Local Similarity 99.7%; Pred. No. 7.6e-286;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 813 GGTACTACTCATAGATTACTGCAACAGTTCCTCCCTGGAACCTGGAAGTTCTTCTGCTG 872
Db 2946 GGAAGAACTCATAGATTACTGCAACAGTTCCTCCCTGGAACCTGGAAGTTCTTCTGCTG 3005
QY 873 GCTTACAGAACTGAAACAACTGCGAATGCTTACAGAGTCTACCGTAAAGGAAAGCT 932
Db 3006 GCTTACAGAACTGAAACAACTGCGAATGCTTACAGAGTCTACCGTAAAGGAAAGCT 3065
QY 933 CCTAGAAGACTCCAAGGAGTAAAGAGTCTGATGAAACAATGGCAAGACCTCCCAAGTGA 992
Db 3066 CCTAGAAGACTCCAAGGAGTAAAGAGTCTGATGAAACAATGGCAAGACCTCCCAAGTGA 3125
QY 993 AATTGAAGCTCACAGAGTGTATCAACCTGGATGAAACAGCCAAACAACTCTGAG 1052
Db 3126 AATTGAAGCTCACAGAGTGTATCAACCTGGATGAAACAGCCAAACAACTCTGAG 3185
QY 1053 ATCCCTGGAAGTTCGGATGATGAGTCTCTTACAAAGACGTTTGGATATCATGAACCT 1112
Db 3186 ATCCCTGGAAGTTCGGATGATGAGTCTCTTACAAAGACGTTTGGATATCATGAACCT 3245
QY 1113 CAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTTAGTCCCATTTTGAAGCCAGTTC 1172
Db 3246 CAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTTAGTCCCATTTTGAAGCCAGTTC 3305
QY 1173 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGTGTGGCTACAGTGA 1232
Db 3306 TGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGTGTGGCTACAGTGA 3365
QY 1233 AGATGATGAATTAAGCGGCGAGCCTATTGAGGCGGACTTTCAGAGCAATTCAGAGCA 1292
Db 3366 AGATGATGAATTAAGCGGCGAGCCTATTGAGGCGGACTTTCAGAGCAATTCAGAGCA 3425
QY 1293 GAACGATGATACATAGGCGCTTCAAGAGGAATTTGAAAACTTAAGAACCTGTATCATGAG 1352
Db 3426 GAACGATGATACATAGGCGCTTCAAGAGGAATTTGAAAACTTAAGAACCTGTATCATGAG 3485
QY 1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGAGCCCTTTGGAAGGACTAGAGAACT 1412
Db 3486 TACTCTTGAGACTGTACGAATATTTCTGACAGAGAGCCCTTTGGAAGGACTAGAGAACT 3545
QY 1413 CTACAGAGCCAGAGAGTGTCTCTGAGGAGAGAGCCAGAAATGTCTACCTGGCTTCT 1472
Db 3546 CTACAGAGCCAGAGAGTGTCTCTGAGGAGAGAGCCAGAAATGTCTACCTGGCTTCT 3605
QY 1473 ACGAAAGCGCTGAGAGGTCATATCTGAGTGGGAAAAATTTGAACCTGCTCCGCTGA 1532
Db 3606 ACGAAAGCGCTGAGAGGTCATATCTGAGTGGGAAAAATTTGAACCTGCTCCGCTGA 3665
QY 1533 CTGCGCAGAGAAAAATAGATGAGAGCCCTTGAAGACTCCAGGAACTTCAAGAGCCACGGA 1592

Db	3666	CTGGCAGAGAAAATAGATGAGACCCCTTGAAGACTCCAGGAATCTCAAGAGGCCACGGA	3725
Qy	1593	TTGAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATCTTGGGAGCCCGTGGG	1652
Db	3726	TTGAGCTGGACCTCAAGCTGCCCAAGCTGAGGTGATCAAGGGATCTTGGGAGCCCGTGGG	3785
Qy	1653	CGATCTCTCATGACTCTCTCCCAAGTCACTCCAGAGAACTCAAGGCACTTCGAGGAGA	1712
Db	3786	CGATCTCTCATGACTCTCTCCCAAGTCACTCCAGAGAACTCAAGGCACTTCGAGGAGA	3845
Qy	1713	AATTGGCCCTCTGAAGAGAACTGAGCCACGTCATCAATGACCTTGTCCGCAAGTTACAC	1772
Db	3846	AATTGGCCCTCTGAAGAGAACTGAGCCACGTCATCAATGACCTTGTCCGCAAGTTACAC	3905
Qy	1773	TTTGGGCATTCAGCTCTCAGGTATTAACCTCAGCACTCTGGAAGACCTGAACACAGATG	1832
Db	3906	TTTGGGCATTCAGCTCTCAGGTATTAACCTCAGCACTCTGGAAGACCTGAACACAGATG	3965
Qy	1833	GAAGCTTCTGCAAGGTGGCCGTGAGGACCGAGTCAGGCAAGCTGCAATGAAGCCCAAGGGA	1892
Db	3966	GAAGCTTCTGCAAGGTGGCCGTGAGGACCGAGTCAGGCAAGCTGCAATGAAGCCCAAGGGA	4025
Qy	1893	CTTTGGTCAGCACTCAGCACTTCTTCCAGCTCTGTCAGGCTCTGTCAGGAGAGGC	1952
Db	4026	CTTTGGTCAGCACTCAGCACTTCTTCCAGCTCTGTCAGGCTCTGTCAGGAGAGGC	4085
Qy	1953	CATCTCGCCAAACAAAGTGGCCCTACTATATCAACACAGAGACTCAAAACA	2001
Db	4086	CATCTCGCCAAACAAAGTGGCCCTACTATATCAACACAGAGACTCAAAACA	4134
RESULT 4			
AX538622			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1. Chamberlain, J.S. and Harper, S.Q.			
Mini-dystrophin nucleic acid and peptide sequences			
Patent: WO 0229056-A 42 11-APR-2002;			
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)			
Location/Qualifiers			
1. 8689			
/organism="synthetic construct"			
/mol_type="genomic DNA"			
/db_xref="taxon:32630"			
/note="synthetic"			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
Qy	813	GGTACTCTACTCATGATTAAGTGAAGAGTCTGATGAAACAAATGGCAAGCTTCCAGTGA	872
Db	2992	GGAAAGAACTCATGATTAAGTGAAGAGTCTGATGAAACAAATGGCAAGCTTCCAGTGA	3051
Qy	873	GCTTACAGAACTGAAACAACTGCAATGCTCTCAGAGAGTCTACCCGTAAGGAAGGCT	932
Db	3052	GCTTACAGAACTGAAACAACTGCAATGCTCTCAGAGAGTCTACCCGTAAGGAAGGCT	3111
Qy	933	CCTAGAAGACTCCAGAGGAGTAAAGAGTCTGATGAAACAAATGGCAAGCTTCCAGTGA	992
Db	3112	CCTAGAAGACTCCAGAGGAGTAAAGAGTCTGATGAAACAAATGGCAAGCTTCCAGTGA	3171
Qy	993	AATTGAAGCTCAGCAGATGTTTATCAACAACTGGATGAACACAGAGACTCAAAACA	1052

RESULT 5
AX538624
LOCUS
DEFINITION
ACCESSION
VERSION

AX538624
Sequence 44 from Patent WO0229056.
AX538624
AX538624.1 GI:25271175

11443 bp DNA linear PAT 23-NOV-2002

[illegible]

QY	1053	ATCCCTGGAAGTTCCGATGATGCAGTCTCTGTTTACAAAGACAGTTTGGATAACATGAACATT	1112
Db	8500	ATCCCTTGGAAAGTTCCGATGATGCAGTCTCTGTTTACAAAGACAGTTTGGATAACATGAACATT	8559
QY	1113	CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGAGCCAGTTC	1172
Db	8560	CAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACATTAGGTCCCATTTGGAGAGCCAGTTC	8619
QY	1173	TGACCAAGTGAAGCGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA	1232
Db	8620	TGACCAAGTGAAGCGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTGGCTACAGCTGAA	8679
QY	1233	AGATGATGAATTAAGCCGGCAGGCACTTAATGGAGGCGACTTTCAGACAGTTCAGNAGCA	1292
Db	8680	AGATGATGAATTAAGCCGGCAGGCACTTAATGGAGGCGACTTTCAGACAGTTCAGNAGCA	8739
QY	1293	GAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTTAAAGAACCTGTAAATCATGAG	1352
Db	8740	GAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAACCTTAAAGAACCTGTAAATCATGAG	8799
QY	1353	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCTTTCGAAAGGACTAGAGAACT	1412
Db	8800	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCGCTTTCGAAAGGACTAGAGAACT	8859
QY	1413	CTACCAAGGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCACAATGTCTACTCGGCTTCT	1472
Db	8860	CTACCAAGGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCACAATGTCTACTCGGCTTCT	8919
QY	1473	ACGAAAGCAGGCTGAGGAGTCAATCTGAGTGGGAAAAATTGAACTGTCACCTCCGCTGA	1532
Db	8920	ACGAAAGCAGGCTGAGGAGTCAATCTGAGTGGGAAAAATTGAACTGTCACCTCCGCTGA	8979
QY	1533	CTGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA	1592
Db	8980	CTGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGA	9039
QY	1593	TGAGCTGGACCTCAAGCTGGGCCAAGCTGAGGTGATCAAGGGATCTCTGGCAGGCCGTGGG	1652
Db	9040	TGAGCTGGACCTCAAGCTGGGCCAAGCTGAGGTGATCAAGGGATCTCTGGCAGGCCGTGGG	9099
QY	1653	CGATCTCTCTATTGACTCTCTCAGAGTACCTCGAGAAAGTCAAGGCACTTCGAGGAGA	1712
Db	9100	CGATCTCTCTATTGACTCTCTCAGAGTACCTCGAGAAAGTCAAGGCACTTCGAGGAGA	9159
QY	1713	AATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC	1772
Db	9160	AATTGGCGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGCTCGCCAGCTTACCAC	9219
QY	1773	TTTGGGCATTGAGTCTACCGTATACCTCAGCACTCTCGAGAGACTGAACACCAGATG	1832
Db	9220	TTTGGGCATTGAGTCTACCGTATACCTCAGCACTCTCGAGAGACTGAACACCAGATG	9279
QY	1833	GAAGCTTCTGAGGTGGCCGTGCGAGGACCGAGTCAAGGAGTCAATGACCTGAAAGCCACAGGGA	1892
Db	9280	GAAGCTTCTGAGGTGGCCGTGCGAGGACCGAGTCAAGGAGTCAATGACCTGAAAGCCACAGGGA	9339
QY	1893	CTTTGGTCCAGCATCTCAGCACTTTCTTTTTCACGCTGTCCAGGGTCCCTGGGAGAGGC	1952
Db	9340	CTTTGGTCCAGCATCTCAGCACTTTCTTTTTCACGCTGTCCAGGGTCCCTGGGAGAGGC	9399
QY	1953	CATCTCGCAACAAAGTGCCTACTATATCAACACAGAGACTCAAAACA	2001
Db	9400	CATCTCGCAACAAAGTGCCTACTATATCAACACAGAGACTCAAAACA	9448
RESULT	7		
HSDMDR			
LOCUS	HSDMDR	12446 bp	mRNA linear PRI 12-SEP-1993
DEFINITION	Human mRNA for dystrophin.		
ACCESSION	X14298		
VERSION	X14298.1	GI:30845	
KEYWORDS	Dmd gene; Duchenne muscular dystrophy; dystrophin.		

[illegible]

RMQEVCKOLSDRKRLEBQXNLSRFDRLNEFLVWLEADNVIASIPLEPGKEOOLKE
KLEOVKLVLEPLPQGLIKQLNETGGPVLVSAPISPEODKLENKILKOTNLOMKVS
RALPEKQGEIAGIKDGLGKEDLEQLNLLWLSPIKQNEJLITQNFQSGPFD
VKETEIAVQKQDVEEILISQOHLYKEKPAQOPVKRKLDESSEKWNRLLOQLRA
KOPDLAPGLTIGASPTQVTLVPTVKETAISKLEMPSLMVEPALADFNRAWT
ELTDLMLSDQVTKQSRVWVGLDINEMIKKATQMDLEQRPQLEBELITAAQNLK
NKTNSOEARTITDRTERIONQWDEVOELHONRQOLNEMKDVQWLEAKEEABOVL
GQARAKLESKSGPYVDALOKKITETKOLAKDLQWOTNDVANDLALKLRLDYAD
DIRKVKHMTENINASHSRHKBVSREALESTHLLQOPDLDEKFLAWLTEABTTA
NVLDQATRKERLLEDKSKVEMKQWQDLQDSEIHAHTDVYHNLDSQKILRSLEGS
DAVILQRLNLMNFKSELKSLNIRSHLEASSQWKRKLHUSLOELLVQLQKDEL
SRQAPITGDFPAVQKQDVRHAFKELTKKEPVIMSTLETVRIFTEPLEGLEKLYQ
EPRELPEPRAQVNTLLRQKAEVNTWEKLNLAEMQKIDETLERLOELQAYQ
ELDLKROAEVIGKSWQPVGDLILDSLODHLKVKALRGEIAPLAKENTSHVNDLAROL
TTLGLOLSPVNLSTLDELNTRKLLQVADVEDRVOLHREHDFGPASOFLSTSVQGP
WESAISPKNVPYVINEHETCTQWDEPKMTELQSLADLNNVPSVATAMKLRLOKA
LCJDLJLSAACDALQOHLNKLQNDQMDILOIINCLITTYIDRLEHNNLNVPLCVD
MCNLWLVYDTRGRIRVLSFKTIIISLCKAHLDEKRYLFKQVASTGFCDDRL
GLLHDSIGITPQGLGEVAFGSGNIEPSVRSQFANNKPEIETALFLDMWRLEPQSM
VMLPVLHRAVAATKHAQAKNICKEPIIGFRYSLKHPYDIQOSCFPSGRVAKGH
KMHVPVEYCTPTTSGEDVRDFAKLNKFRKRVFAKHPRMGYLPVQTVLEGNMET
PVTILNFWPVDSPASSPOLSHDDTHSRIEHVASLAEMENSGVYNDISPNSID
DEHLLQHYCOSLNQSPQSPPSQAQILISLESEERGELERILADLEENENLOAEY
DRUKQEHKGLRLLSPPEMPTSPQSPRDAELIAEAKLLQHKGRLEARNMOILEDH
NKQESQLHKLRLQLEPQKQAEKVNQVTSSTLSQSDSPQSMRLRVVGSQTSDSM
GEEDLLSPPODTSTGLEEVNEQLNNSFPSSRGHNTPGKPMREBDM"

BASE COUNT 4135 a 2524 c 2876 g 2911 t
ORIGIN

Query Match 59.1%; Score 1182.6; DB 9; Length 12446;

Best Local Similarity 99.7%; Pred. No. 7.9e-286; Indels 0; Gaps 0;

Matches 1185; Conservative 0; Mismatches 4;

813 GGTACCTACTCATAGATTACTGCAACAGTTTCCCTCGGACCTCGAAGGTTTCTTGCCCTG 872

8150 GGAAGAACTCATAGATTACTGCAACAGTTTCCCTCGGACCTCGAAGGTTTCTTGCCCTG 8209

873 GCTTACAGAGCTGAACAACCTGCGAATGCTTACAGGATGCTACCGGTAAAGAAAGGCT 932

8210 GCTTACAGAGCTGAACAACCTGCGAATGCTTACAGGATGCTACCGGTAAAGAAAGGCT 8269

933 CCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAACAATGGAAGACCTCCCAAGGTGA 992

8270 CCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAACAATGGAAGACCTCCCAAGGTGA 8329

993 AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAGAAACAGCCAAATCTCTGAG 1052

8330 AATTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAGAAACAGCCAAATCTCTGAG 8389

1053 ATCCCTGGAAGTTCCGATGATGAGTCCCTGTTTACAAAGAGCTTTGATACATGAATCTT 1112

8390 ATCCCTGGAAGTTCCGATGATGAGTCCCTGTTTACAAAGAGCTTTGATACATGAATCTT 8449

1113 CAAGTGGAGTGAATCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTTC 1172

8450 CAAGTGGAGTGAATCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTTC 8509

1173 TGACAGTGGAGGCTGTGACCTTTCTCTGACGAGACTTCTGGTGGCTACAGCTGA 1232

8510 TGACAGTGGAGGCTGTGACCTTTCTCTGACGAGACTTCTGGTGGCTACAGCTGA 8569

1233 AGATTGATGAATTAAGCCGGCAGGACCTATTGAGGAGGAGCTTTCCAGCAGTTTCAAGAGCA 1292

8570 AGATTGATGAATTAAGCCGGCAGGACCTATTGAGGAGGAGCTTTCCAGCAGTTTCAAGAGCA 8629

1293 GAAAGATGATCATAGGCTTTCAAGAGGAAATTGAAGAACTTAAGAACTTATCATGAG 1352

8630 GAAAGATGATCATAGGCTTTCAAGAGGAAATTGAAGAACTTAAGAACTTATCATGAG 8689

1353 TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCTTTTGAAGGACTTAGAGAACT 1412

8690 TACTCTTGAGACTGTACGAATATTCTTGACAGAGCAGCTTTTGAAGGACTTAGAGAACT 8749

QY 1413 CTACAGAGCCAGAGAGCTGCTCTCTGAGAGAGAGCCAGAAATGTCACTCGGCTTCT 1472

Db 8750 CTACAGAGCCAGAGAGCTGCTCTCTGAGAGAGAGCCAGAAATGTCACTCGGCTTCT 8809

QY 1473 ACAGAAACAGCTGAGAGGTCAATCTAGTGGGAAAAATTGAACCTGCACTCGGCTGA 1532

Db 8810 ACAGAAACAGCTGAGAGGTCAATCTAGTGGGAAAAATTGAACCTGCACTCGGCTGA 8869

QY 1533 CTGCGACAGAAAAATAGATGAGACCTTGAAGAACTTCAAGAGGCACTTCAAGAGGCACTGA 1592

Db 8870 CTGCGACAGAAAAATAGATGAGACCTTGAAGAACTTCAAGAGGCACTTCAAGAGGCACTGA 8929

QY 1593 TGAGCTGGAACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGAGGATCTCTGGGAGCCCTGGG 1652

Db 8930 TGAGCTGGAACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGAGGATCTCTGGGAGCCCTGGG 8989

QY 1653 CGATCTCTCATTTGACCTCTCTCAAGATCACTCGAGAAAGTCAAGGCACTTTCAGAGAGA 1712

Db 8990 CGATCTCTCATTTGACCTCTCTCAAGATCACTCGAGAAAGTCAAGGCACTTTCAGAGAGA 9049

QY 1713 AATTGCGCCTCTGAAAGAGAAAGCTGAGCCACGTCATATGACCTTGTCTCGCAGCTTACAC 1772

Db 9050 AATTGCGCCTCTGAAAGAGAAAGCTGAGCCACGTCATATGACCTTGTCTCGCAGCTTACAC 9109

QY 1773 TTTGGGCACTTCAAGCTTCAAGCTGATTAACCTTCAAGACCTTCAAGACCTTCAAGACCTG 1832

Db 9110 TTTGGGCACTTCAAGCTTCAAGCTGATTAACCTTCAAGACCTTCAAGACCTTCAAGACCTG 9169

QY 1833 GAAGCTTCTGAGCTGGCCCTCGAGGACCGAGTCAAGCTGAGCTGATCAAGCCACACAGGA 1892

Db 9170 GAAGCTTCTGAGCTGGCCCTCGAGGACCGAGTCAAGCTGAGCTGATCAAGCCACACAGGA 9229

QY 1893 CTTTGGTCCAGATCTCAGACATTTCTTCCAGCTCTGTCAGGCTCCCTGGGAGAGAGC 1952

Db 9230 CTTTGGTCCAGATCTCAGACATTTCTTCCAGCTCTGTCAGGCTCCCTGGGAGAGAGC 9289

QY 1953 CATCTCGGCAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAACA 2001

Db 9290 CATCTCGGCAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAACA 9338

RESULT 8

AX409637 13957 bp DNA linear PAT 14-JUN-2002

LOCUS AX409637 Sequence 2284 from Patent WO0229103.

DEFINITION AX409637

ACCESSION AX409637

VERSION AX409637.1 GI:21442342

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Alves, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.

TITLE Gene expression profiles in liver cancer

JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;

GENE LOGIC INC (US)

FEATURES

Location/Qualifiers

1. 13957

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/note="EMBL/GenBank Accession No. M18533"

BASE COUNT 4602 a 2781 c 3122 g 3452 t

ORIGIN

Query Match 59.1%; Score 1182.6; DB 6; Length 13957;

Best Local Similarity 99.7%; Pred. No. 7.9e-286; Indels 0; Gaps 0;

Matches 1185; Conservative 0; Mismatches 4;

813 GGTACCTACTCATAGATTACTGCAACAGTTTCCCTCGGACCTCGAAGGTTTCTTGCCCTG 872

8260 GGAAGAACTCATAGATTACTGCAACAGTTTCCCTCGGACCTCGAAGGTTTCTTGCCCTG 8319

QY	873	GCTTACAGAGCTGAACCAACTGCCAATGTCCTACAGATGCTACCGTGAAGGAAGGCT	932
Db	8320	GCTTACAGAGCTGAACCAACTGCCAATGTCCTACAGATGCTACCGTGAAGGAAGGCT	8379
QY	933	CTAGAGAGCTCCAGGAGTAAGAGCTGATGAACAATGGCAAGACCTCCAAAGTGA	992
Db	8380	CTAGAGAGCTCCAGGAGTAAGAGCTGATGAACAATGGCAAGACCTCCAAAGTGA	8439
QY	993	AATTGAAGCTCACAGAGTGTATATCAAACTGGATGAACCAAGCCAAATAATCTCTGAG	1052
Db	8440	AATTGAAGCTCACAGAGTGTATATCAAACTGGATGAACCAAGCCAAATAATCTCTGAG	8499
QY	1053	ATCCCTGGAGGTTCCGATGATGCTCTCTCAAGAGAGCTTTGGATACATGAATCT	1112
Db	8500	ATCCCTGGAGGTTCCGATGATGCTCTCTCAAGAGAGCTTTGGATACATGAATCT	8559
QY	1113	CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCAAGTTC	1172
Db	8560	CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCAAGTTC	8619
QY	1173	TCACAGTGGAGGCTCTGCACCTTTCTCTGAGGAGCTTTCCAGAGCTTTTCAGAGCA	1232
Db	8620	TCACAGTGGAGGCTCTGCACCTTTCTCTGAGGAGCTTTCCAGAGCTTTTCAGAGCA	8679
QY	1233	AGATGATGAATTAAGCCGAGGAGCTTATTCGAGGAGCTTTCCAGAGCTTTTCAGAGCA	1292
Db	8680	AGATGATGAATTAAGCCGAGGAGCTTATTCGAGGAGCTTTCCAGAGCTTTTCAGAGCA	8739
QY	1293	GAAAGTGTACATAGGAGCTTTCAAGAGGAGTGAATACTAAAGAACTGTATATCATGAG	1352
Db	8740	GAAAGTGTACATAGGAGCTTTCAAGAGGAGTGAATACTAAAGAACTGTATATCATGAG	8799
QY	1353	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGCACTAGAGAACT	1412
Db	8800	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGCACTAGAGAACT	8859
Query Match 59.1%; Score 1182.6; DB 6; Length 13957;			
Best Local Similarity 99.7%; Pred. No. 7.9e-286;			
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	813	GCTACCTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAGCTTTCTTGCCTG	872
Db	8260	GGAAGAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAGCTTTCTTGCCTG	8319
QY	873	GCTTACAGAGAGCTGAACCAACTGCCAATGTCCTACAGAGTGTACCCGTGAAGAAAGCT	932
Db	8320	GCTTACAGAGAGCTGAACCAACTGCCAATGTCCTACAGAGTGTACCCGTGAAGAAAGCT	8379
QY	933	CTTGAAGAAGCTCCAAAGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAAGTGA	992
Db	8380	CTTGAAGAAGCTCCAAAGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCCAAAGTGA	8439
QY	993	AATTGAAGCTCACAGAGTGTATATCAAACTGGATGAACCAAGCCAAATAATCTCTGAG	1052
Db	8440	AATTGAAGCTCACAGAGTGTATATCAAACTGGATGAACCAAGCCAAATAATCTCTGAG	8499
QY	1053	ATCCCTGGAGGTTCCGATGATGCTCTCTCAAGAGAGCTTTGGATACATGAATCT	1112
Db	8500	ATCCCTGGAGGTTCCGATGATGCTCTCTCAAGAGAGCTTTGGATACATGAATCT	8559
QY	1113	CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCAAGTTC	1172
Db	8560	CAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAAGTCCCATTTGGAAGCAAGTTC	8619
QY	1173	TCACAGTGGAGGCTCTGCACCTTTCTCTGAGGAGCTTTCCAGAGCTTTTCAGAGCA	1232
Db	8620	TCACAGTGGAGGCTCTGCACCTTTCTCTGAGGAGCTTTCCAGAGCTTTTCAGAGCA	8679
QY	1233	AGATGATGAATTAAGCCGAGGAGCTTATTCGAGGAGCTTTCCAGAGCTTTTCAGAGCA	1292
Db	8680	AGATGATGAATTAAGCCGAGGAGCTTATTCGAGGAGCTTTCCAGAGCTTTTCAGAGCA	8739
QY	1293	GAAAGTGTACATAGGAGCTTTCAAGAGGAGTGAATACTAAAGAACTGTATATCATGAG	1352
Db	8740	GAAAGTGTACATAGGAGCTTTCAAGAGGAGTGAATACTAAAGAACTGTATATCATGAG	8799
QY	1353	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGCACTAGAGAACT	1412
Db	8800	TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGCACTAGAGAACT	8859

NVLQDATERKRLLEDSDSGVXELMKQWDLQGEIEAHYDTHVHNDENSKILBSLEGSD
DAVLLQRLDNPNFNSBELRRKSLNIRSHLEASDDQWRLHLSIQELLVWQKQDEL
SROAPIGDGDFAVQKQNDVHRAFKRELTKEPVMISLETVRIFLFEQPLEGLEKUIQ
EPRELPEERQNTVRLLRQAEBENTWEKMLNHSADWQKRIDETLERLQELQBATD
ETDLQKQAEVYKGSWOPVGLLDSQDHLKVKALRGETAPLKENVSHVNDLBAROL
FLTGIOQSPVNLSTDLNTRWKLQVAVDRVQLHEARDPQASQVHFLSTSVQGP
WERASINKVPIYINHEITQTCWDPKPMWLSYSLADLNNRFSAYRTAMKRLRLOKA
LCLDLLSUACADALDQNDQNDPMDIILQINCLTITLYDRLEQSHNNLVNPLVCD
MCLMWLNVDYGTGRIRVLVSFKTIGIISLCAHLEKDKYRLFKQVASTGFCDOQRLL
GLLLHDSIQIPROQLGEVASFQSNIEPVSRSCTFOFANNKEIEEALFLDWRLEPQSM
VWLPLVHLVAAETAKHQAQKNIKECPIIGFRYSRLKFNHYDICSQFFSFGVAKHG
KWHVMEVYCTPTTSGEDVRFAKLVKFKRTKRYFAKHPRMGVLYPQTVLEGDNMET
PVTILNFWYVDSAPASSPOLSDHDTTHRIEYHVASRLAEMENSGVLYNDISNESID
DEHLLIOHYCSLQNDSPSPSPAPOLISLESEERLELILADLEENRMLQABY
DRLKQKHGLSLPSPPEMPTSPSPADAEILIAEKLLROHKGLEARMOLLEDEH
NQLESHLRILRLQLEPOAEKVNQNTVSPFSLRSRQSSQPMILRVVGSQTSDSM
GEEDLSPQDTSSTGLEVMQSLNPFSSRGRNTPGPKMKEDTM
BASE COUNT 4602 a 2781 c 3122 g 3452 t

Query Match 59.1%; Score 1182.6; DB 9; Length 13957;
Best Local Similarity 99.7%; Pred. No. 7.9e-286;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCGTG 872
DB 8260 GGAAGAACTCATAGTTACTCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCGTG 8319
QY 873 GCTTACAGAGCTGAAGAGCTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCGTG 932
DB 8320 GCTTACAGAGCTGAAGAGCTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCGTG 8379
QY 933 CTTAGAGACTCCAAGGAGTAAAGAGCTGATGAACATGGCAAGACCTCCCAAGTGA 992
DB 8380 CTTAGAGACTCCAAGGAGTAAAGAGCTGATGAACATGGCAAGACCTCCCAAGTGA 8439
QY 993 AATTGAAGCTCACAGATGTTTATCAACCTGATGAAACAGCCAAACAAATTCCTGAG 1052
DB 8440 AATTGAAGCTCACAGATGTTTATCAACCTGATGAAACAGCCAAACAAATTCCTGAG 8499
QY 1053 ATCCCTGGAAAGTTCCGATGATGCACTCTGTTACAAAGAGCTTTGGATACATGAACCT 1112

DB 8980 CTGGCAGAGAAAAATAGATGAGACCTTTGAAGAGACTCCAGGAACCTTCAAGAGGCCACGGA 9039
QY 1593 TGAGCTGGACCTCAGCTGCGCCAGCTGAGTGATCAAGGGATCCTCGCAGCCCGTGGG 1652
DB 9040 TGAGCTGGACCTCAGCTGCGCCAGCTGAGTGATCAAGGGATCCTCGCAGCCCGTGGG 9099
QY 1653 CGATCTCCTCATTTGATCTCTCTCCAAAGATCACTCGAGAAAGTCAAGGACCTTCGAGGGA 1712
DB 9100 CGATCTCCTCATTTGATCTCTCTCCAAAGATCACTCGAGAAAGTCAAGGACCTTCGAGGGA 9159
QY 1713 AATTGGCTCTGAAAGAGAGCTGAGCCAGCTCAATCACTGCTCGCCAGCTTACCAC 1772
DB 9160 AATTGGCTCTGAAAGAGAGCTGAGCCAGCTCAATCACTGCTCGCCAGCTTACCAC 9219
QY 1773 TTTGGCAATTCAGCTCTCACCCTTAATCACTCAGACTCTGGAAGACCTTGAACACCATG 1832
DB 9220 TTTGGCAATTCAGCTCTCACCCTTAATCACTCAGACTCTGGAAGACCTTGAACACCATG 9279
QY 1833 GAAGCTTCTGAGTGGCGCTCGAGGACCGAGTCAAGGAGCTGCATGAAGCCACACAGGA 1892
DB 9280 GAAGCTTCTGAGTGGCGCTCGAGGACCGAGTCAAGGAGCTGCATGAAGCCACACAGGA 9339
QY 1893 CTTTGGTCCAGATCTCAGCACTTTCTTTCCACGTTGTCAGGGTCCCTGGAGAGAGC 1952
DB 9340 CTTTGGTCCAGATCTCAGCACTTTCTTTCCACGTTGTCAGGGTCCCTGGAGAGAGC 9399
QY 1953 CATCTCGCAACAAAGTGCCTTACTATATCAACACGAGACTCAAA 2001
DB 9400 CATCTCGCAACAAAGTGCCTTACTATATCAACACGAGACTCAAA 9448

RESULT 11
AR220819 13977 bp DNA linear PAT 26-SEP-2002
LOCUS Sequence 60 from patent US 6426186.
DEFINITION AR220819
ACCESSION AR220819.1 GI:23327696
VERSION
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 13977)
AUTHORS Jones, K.A., Volkmut, W. and Walker, M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES Location/Qualifiers
source 1..13977
BASE COUNT 4596 a 2765 c 3120 g 3453 t 43 others
ORIGIN

Query Match 59.1%; Score 1182.6; DB 6; Length 13977;
Best Local Similarity 99.7%; Pred. No. 7.9e-286;
Matches 1185; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 813 GGTACCTACTCATAGATTACTCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCGTG 872
DB 8260 GGAAGAACTCATAGTTACTCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCGTG 8319
QY 873 GCTTACAGAGCTGAAGAGCTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCGTG 932
DB 8320 GCTTACAGAGCTGAAGAGCTGCAACAGTTCCCTCGACCTGGAAAAGTTTCTTGCCGTG 8379
QY 933 CTTAGAGACTCCAAGGAGTAAAGAGCTGATGAACATGGCAAGACCTCCCAAGTGA 992
DB 8380 CTTAGAGACTCCAAGGAGTAAAGAGCTGATGAACATGGCAAGACCTCCCAAGTGA 8439
QY 993 AATTGAAGCTCACAGATGTTTATCAACCTGATGAAACAGCCAAACAAATTCCTGAG 1052
DB 8440 AATTGAAGCTCACAGATGTTTATCAACCTGATGAAACAGCCAAACAAATTCCTGAG 8499
QY 1053 ATCCCTGGAAAGTTCCGATGATGCACTCTGTTACAAAGAGCTTTGGATACATGAACCT 1112

8500 ATCCCTGGAGGTTCCGATGATCGAGTCTGTTACAGAGCGTTTGGATAACATGAACCTT 8559
1113 CAACTGGAGTGAATTCGGAAGAAAGTCTCTCAACATAGTCTCCATTTGGAGCCAGTTC 1172
8560 CAACTGGAGTGAATTCGGAAGAAAGTCTCTCAACATAGTCTCCATTTGGAGCCAGTTC 8619
1173 TGACCACTGGAGGCTCTGACCTTTCTCTGAGGAACTTCTGCTGTGGCTACAGCTGAA 1232
8620 TGACCACTGGAGGCTCTGACCTTTCTCTGAGGAACTTCTGCTGTGGCTACAGCTGAA 8679
1233 AGATGATGAATTAAGCCGAGGACCTATTTGAGGCGGCTTTCCAGAGCTTCAAGAGA 1292
8680 AGATGATGAATTAAGCCGAGGACCTATTTGAGGCGGCTTTCCAGAGCTTCAAGAGA 8739
1293 GAACGATGATACATAGGCTCTCAAGAGGAAATTTGAAGAACTTAAGAACTTGAATCATGAG 1352
8740 GAACGATGATACATAGGCTCTCAAGAGGAAATTTGAAGAACTTGAATCATGAG 8799
1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAGGACTAGAGAACT 1412
8800 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAGGACTAGAGAACT 8859
1413 CTACAGAGGCCAGAGAGCTGCTCTCTGAGGAGAGCCGAGATGTCACTCGGCTTCT 1472
8860 CTACAGAGGCCAGAGAGCTGCTCTCTGAGGAGAGCCGAGATGTCACTCGGCTTCT 8919
1473 ACAGAAAGCAGCTTGAGGAGTCAATCTAGTGGGAAATTTGAACCTGCACTCCGCTGA 1532
8920 ACAGAAAGCAGCTTGAGGAGTCAATCTAGTGGGAAATTTGAACCTGCACTCCGCTGA 8979
1533 CTGGCAGAGAAATAGATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGCCACGGA 1592
8980 CTGGCAGAGAAATAGATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGCCACGGA 9039
1593 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGATCAAGGATCTCTGAGCCGCTGGG 1652
9040 TGAGCTGGACCTCAAGCTGCGCAAGCTGAGGATCAAGGATCTCTGAGCCGCTGGG 9099
1653 CGATCTCTCTCAATGACTCTCTCAAGATACCTCGAGAAAGTCAAGGCACTTCCAGGAGA 1712
9100 CGATCTCTCTCAATGACTCTCTCAAGATACCTCGAGAAAGTCAAGGCACTTCCAGGAGA 9159
1713 AATTCGCTCTGAAAGAGAGCTGAGCCAGCTCAATGACTCTGCTGCGCAGCTTACCAC 1772
9160 AATTCGCTCTGAAAGAGAGCTGAGCCAGCTCAATGACTCTGCTGCGCAGCTTACCAC 9219
1773 TTTGGGCAATTCAGCTCTCAAGCTATACCTCAGCACTCTGGAAGACCTGAACACCAAGATG 1832
9220 TTTGGGCAATTCAGCTCTCAAGCTATACCTCAGCACTCTGGAAGACCTGAACACCAAGATG 9279
1833 GAAGCTTCTGAGTGGCGCTGAGGAGCCAGTCAAGGAGCTGATGAGCCACAGGGA 1892
9280 GAAGCTTCTGAGTGGCGCTGAGGAGCCAGTCAAGGAGCTGATGAGCCACAGGGA 9339
1893 CTTTGGTCCAGACTCTCAGCACTTTCTTCTGAGTCTGCTGAGGCTCTTGGAGAGAGC 1952
9340 CTTTGGTCCAGACTCTCAGCACTTTCTTCTGAGTCTGCTGAGGCTCTTGGAGAGAGC 9399
1953 CATCTGCCAAACAAAGTGGCCCTACTATATCAACAGAGACTCAACA 2001
9400 CATCTGCCAAACAAAGTGGCCCTACTATATCAACAGAGACTCAACA 9448

RESULT 12
LOCUS AX538620 5339 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM

artificial sequences.
1 Chamberlain, J.S. and Harper, S.Q.
Mini-dystrophin nucleic acid and peptide sequences
Patent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..5339
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1638 a 1191 c 1187 g 1323 t
ORIGIN
Query Match 59.0%; Score 1180; DB 6; Length 5339;
Best Local Similarity 100.0%; Pred. No. 3.4e-285;
Matches 1180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 822 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAGCTTTCTTGCCTGGCTTACAGA 881
DB 1546 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAGCTTTCTTGCCTGGCTTACAGA 1605
QY 882 AGCTGAAACAACTGCCAATGCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGA 941
DB 1606 AGCTGAAACAACTGCCAATGCTCTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAGA 1665
QY 942 CTCGAAGGAGTAAAGAGCTGATGAACAAATGCGCAGACCTCCAGGTGAAATTTGAAGC 1001
DB 1666 CTCGAAGGAGTAAAGAGCTGATGAACAAATGCGCAGACCTCCAGGTGAAATTTGAAGC 1725
QY 1002 TCACACAGATGTTTATCAACAACTGGATGAAACAGCCAAAGAAATCTCTGAGTCCCTGGA 1061
DB 1726 TCACACAGATGTTTATCAACAACTGGATGAAACAGCCAAAGAAATCTCTGAGTCCCTGGA 1785
QY 1062 AGGTTCGATGATGCGATCTCTGTTACAAAGCGTTTGGATTAACATGAACCTTCAAGTGGAG 1121
DB 1786 AGGTTCGATGATGCGATCTCTGTTACAAAGCGTTTGGATTAACATGAACCTTCAAGTGGAG 1845
QY 1122 TGAATTCGGAAGAAAGTCTCTCAACATTTAGTCCATTTGGAAGCCAGTTCCTGACCAAGT 1181
DB 1846 TGAATTCGGAAGAAAGTCTCTCAACATTTAGTCCATTTGGAAGCCAGTTCCTGACCAAGT 1905
QY 1182 GAAGCGTCTGACACTTTCTCTGAGGAACTTCTGGTGTGGCTACAGCTGAAGATGATGA 1241
DB 1906 GAAGCGTCTGACACTTTCTCTGAGGAACTTCTGGTGTGGCTACAGCTGAAGATGATGA 1965
QY 1242 ATTAAGCCGAGGACCTATTTGGAGCGACTTTCCAGCAGTTCAGAGAGCAAGACGATGT 1301
DB 1966 ATTAAGCCGAGGACCTATTTGGAGCGACTTTCCAGCAGTTCAGAGAGCAAGACGATGT 2025
QY 1302 ACATAGGCGCTTCAAGAGGGAATTTGAAACCTTAAAGAACTTAAAGAACTTAAAGAACTTAA 1361
DB 2026 ACATAGGCGCTTCAAGAGGGAATTTGAAACCTTAAAGAACTTAAAGAACTTAAAGAACTTAA 2085
QY 1362 GACTGTACGAATATTTCTGACAGAGCGCTTTTGGAGGACTAGAGAACTTCTACCAAGGA 1421
DB 2086 GACTGTACGAATATTTCTGACAGAGCGCTTTTGGAGGACTAGAGAACTTCTACCAAGGA 2145
QY 1422 GCCCAGAGAGCTGCTCTCTGAGGAGAGCCAGAACTGCTCACTCGGCTTCTACCAAGGA 1481
DB 2146 GCCCAGAGAGCTGCTCTCTGAGGAGAGCCAGAACTGCTCACTCGGCTTCTACCAAGGA 2205
QY 1482 GGCTGAGGAGTCAATCTGAGTGGGAAATTTGAACCTTCACTCGGCTTCTGAGTGGCAGAG 1541
DB 2206 GGCTGAGGAGTCAATCTGAGTGGGAAATTTGAACCTTCACTCGGCTTCTGAGTGGCAGAG 2265
QY 1542 AAAAATAGATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGGCCACCGGATGAGCTGGA 1601
DB 2266 AAAAATAGATGAGACCTTTGAAAGACTCCAGAACTTCAAGAGGCCACCGGATGAGCTGGA 2325
QY 1602 CCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCTTGGCAGGCCCTGGGGGATCTCT 1661

```
Db 2326 COTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCTGGCGGATCTCTT 2385
Qy 1662 CATTGACTCTCTCCAAAGATCACTCGAGAAAGTCAAGGCATCTCGAGGAGAAATTTGGCC 1721
Db 2386 CATTGACTCTCTCCAAAGATCACTCGAGAAAGTCAAGGCATCTCGAGGAGAAATTTGGCC 2445
Qy 1722 TCTGAAAGAGAAAGTGAAGGACGAGTCAATGACTCTGCTCGCAGCTTACCCTTTGGGCAT 1781
Db 2446 TCTGAAAGAGAAAGTGAAGGACGAGTCAATGACTCTGCTCGCAGCTTACCCTTTGGGCAT 2505
Qy 1782 TCAGTCTCTCACCGTATAACCTCAGACACTCTGAGACACTGAGACCTGAACACGAGATGAAGCTTCT 1841
Db 2506 TCAGTCTCTCACCGTATAACCTCAGACACTCTGAGACACTGAGACCTGAACACGAGATGAAGCTTCT 2565
Qy 1842 GCAGTGGCCGTCGAGAGCCGAGTCAAGGACGAGTCAATGACTCTGAGACCTGAACACGAGATGAAGCTTCT 1901
Db 2566 GCAGTGGCCGTCGAGAGCCGAGTCAAGGACGAGTCAATGACTCTGAGACCTGAACACGAGATGAAGCTTCT 2625
Qy 1902 AGCATCTCAGACTTCTTCTTCCAGCTCTGCTCAGGAGTCCCTGGGAGAGAGCATCTGCC 1961
Db 2626 AGCATCTCAGACTTCTTCTTCCAGCTCTGCTCAGGAGTCCCTGGGAGAGAGCATCTGCC 2685
Qy 1962 AAACAAGTGCCTACTATATCAACACGAGACTCAACA 2001
Db 2686 AAACAAGTGCCTACTATATCAACACGAGACTCAACA 2725

RESULT 13
AF070485 13887 bp mRNA linear MAN 09-DEC-1998
LOCUS Canis familiaris dystrophin mRNA, complete cds.
DEFINITION AF070485
ACCESSION AF070485
VERSION AF070485.1 GI:3982750
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 13887)
AUTHORS Carville, K.S., Mann, C.J., Schatzberg, S.J. and Wilton, S.D.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-1998) ANRI, Pathology, University of Western Australia, Verdun Street, Nedlands, WA 6018, Australia
FEATURES
source 1..13887
/mol_type="mRNA"
/strain="breed Golden retriever"
/db_xref="taxon:9615"
<1..289
290..11332
/codon_start=1
/product="dystrophin"
/protein_id="AAC83646.1"
/db_xref="GI:3982751"
/translation="MLAWEEVEDCYREDVOKXTFTKVNNAQPSKFKGHIENIFSDL
QDGRLLDLLEGTLGQKLPKELKTRVHALNKNVKNALRVKKNVLDVNGISDIVDG
NKLTLGLNIIHLHWQVKNVKNINAGLQCTSEKILLSWQVSRNYPQVNVNFT
TSWSDGALNALIHSHRPDLDFMNSVVCQSQATQRLHFAHNIAXYQIGLEKLDPEY
ATTYPDKSLIMYITSLFVLPOQVYSIAEIOEVMUPRPSQVTRREHFIQHMHYSQ
QITVSLAQGYERAPFPKPRPKSYATQAAVYTTSDTRSPSPHQHLETPEDKSFGRS
LTETANLDSYQALIEVLSWLLSAEDALQAQGEISNDVEEVEKQHPHTEGYMDLTS
HGRVGNVLGMSLIGTKVLSDEEVEQVOMNLSRWECRLVASMESKQSLHKLVL
MDLQNGVLELNDMLTTEETRTKEKEPLGPDIEDLKROVQKHQKLEQVLEQVRL
NKLTHMVVVDESSGDHATALEQLKVLGDRVANI CRWTEDEWLLQDILLKQWRT
EEQCLFSAWLEKEDAVNKHITGFKQCEVLSNQLKVLDTLEKKQTMKQLSL
NQDLISALNTKVAHMEAWNDFAQRWNLVQLEKSSAQISOAVTITPSTQITVY
NETVMTVTRHIIYKHAQELPPPPQKQEIIVDEIRDLVDITELHSMITRSE
AVLOSPEPAIYKKEGNFSDLKKEKVNIAEREAERFKLQDASRAQALVEQMVNRYN
ADSIKQASEQLNSRIWEPICOLLSERLNWLVEYONNIITFYNQLQLEQMTTAENWLT
QPTTTSBTAIKSOLIKCKDIRNLSALQPOIERLKIQSIALKEKGGMFLDFAV
FTNHNQVADVAQAREKELQITIPSLPMRYQETWSTILWIOQSEKLSIPVTVTE
YDIMEQRIGELQALQSSLOEQQNGNLNLTITVKNMSKAPLSLDSIRKYSQEFIEGR
```

```
WKLSSQLVHCQKLEQMAKLRKIQNHITKTKWITVDFVFLKEEWPAALGDSLEIKR
QKCRLLVNDIQTIPSLNVEGAAQCKNAEPEFAGRLTELELRELTNDVYNCRO
VYARKEALGLDITQPSLVNLOKDLSEHMTQAEZEYLEFETKTPDLSQTAVEEMKR
AKEEAQKEAKVKLLTESVNIACAPPAQEAALKDELDTTNTYQMLCTINGRCKT
LEWVWAKHELLSVLEKANKVLSVEVKLTETENISGABEIAEVLDSELENMOHSED
NPNOIRIAQTLTDGVDNDELINLEETFNPSRWELEHAAVRROKLSQSOIAQOEIE
KSLHILQESLSDIQAQKLDQVSMKFLFOKPFANFQKQESKILDEVRKHLPALETKS
ETAQVLSQSIDVAQKLDQVSMKFLFOKPFANFQKQESKILDEVRKHLPALETKS
VQEVVQQLHCVNLKSLSEVSEVEMVITGRIVQKQOTENKELDERVETALK
HNYELGAVTERKQKLEKCLSRMRKEMNALTEWLAATDMELTKRSAVEGMPENLD
SEVWAKTAKEIKOKVHLKSVTEVGEALTVLKGKEMLVEDKSLINSLNIAVTSR
AEWNLILLEVCQKMETFOQNDVITWIIQADALLDESEKKPOCKDILKRAEM
NDIRKQSDRDQAANLANMGDCHRCVPEKISLNHFAAISRIKTGKASIPKE
LEQFNSDKLLEPLEAETIQQVNLKEEDFNKMSDEGTVKELQDGNLQORITD
ERKEEIIKQQLLKNALKDLSRQKKALEISHOWYQYKQADADLLKCLDIEK
KLASPEPRDERKIEIDRELQKKBELNVRRAAGLSEDEGAANVAEFTQQLSKRW
REIEGKFAFRRLNPAQIHTVHEESVAMTEDEPLEISYVSTYLTETHTVSQALSEV
ELLNAPDLCAQDFEDLFKQESLXNIDKSLQISGRIDIIHNKKTALHSATPAERA
KLEALSLDFOFERNVNYKDRQGRFDRSVEMRRFHYDMKILNQWTEAQLFKKT
QIPENWEHAKYKYLKELDGIGQSVVULNATGEBIIQOSSXTASIIQEKLSL
NLRWEVCQKLAERKRLKEEQNIISEBQDVNBEVLMLEADNVANTPLEPQNEQOL
KEKLEQKLABELPQGLILQKLNETGTVLVSAPLSPEDQKLENLKOTNLOMK
VSRNLPKEQEEIEAHVDKLGQLEQNLHLLWLSPIRNLQETIYNQPGFDIKELI
VAVQAQPDVEGILSKGQHYKEKATOPAKRKLJEDLSDDKVVTLQJELRAKQFPG
APGLTTVPAPPSQTVTLTPAVTKETATISLEMPSSLLLEVPALADFNRAVETLTDW
LSLIDRVIKSQSVMGDLDEINEMIIKOKATLDLQORRPOLEELITANQNLKNTSN
QEARITIDRIEIQSDWEQHLQNLQETLQEMKQSTOWLEAKAEAEQVLGQARA
KLESWEKAPYTVDAIKKITEKQKLAQLQRLQVQNDVANDLALALLRDYSADDTKRV
HMITENINASWASIHKRLSEREALEETHRLLQQLPDLKEFLAMEATETANVLQD
ATKRLDLSDSGVRLKMKQWQDLOGESTEAHTDIYHNDLDEQKVLRSLEGSDAALL
QHRDLNNFQGBELRKSNIISHLSEASDQWRLHLSLQELLVQLQKLDLSLQAP
IGGDPFAVQKONDVHFAKRELKTEPVIMSTLETVRIFLTEQPLEGKLEQEPREL
PPEERANVTLLRLKQAEVNTQWEKLVNSADQWKIDLEALERLQELQEADELDLQ
LROAEVIKSQWOPVGLLIDSLQDHLKELKALRGITPLKENVYVNDIAROLITLGI
QLSPYINLTLEDLNTRWKLLQVADIEDRIQLHEAHDGCPASCHFLSTSVQGPWEAI
SPKNPYINHTQTTCWDHPKTELYQSLADLNVRFSAYRTAKRLQKALCIDL
LSLSACDADLQNLKONDQMDILQVINCLITIIDRLQEHNLNVLPLQVDMCLNH
LLNVYDITGRIVLISFTGIIISLCKAHLKRYLFKQVASTGFCQORLGLLHL
DSIQIPROLGEVAFSGSNICEPIIGFVRSLSKFNHYDIQCQFFSGVAKGHMPLV
LHRVAAETAKHOAKNICKEPIIGFVRSLSKFNHYDIQCQFFSGVAKGHMPLV
MVEYCTTSGEDVRDPAKLVKKNPKRYFAKHPRMGYLPVQTVLEGNNMETFVLI
NFWVSAPASSQLSHDPTSHIEHVASLKNWNSGYLNDISI SPNESIDIEHL
IQHYRSLNPSQSPRSPRPAQILISLESEGELELILADLEGRNRLQAEYDRILQK
QHEHGLSLPLSPPEXMPPTSQPSRDAELIAEAKLRQHKGRLEARNQAEYDRILQK
SOLHRLQLLEQPAQAEKAVNGTTVSPSTSLQSSDSQPMLLRLRVGSGISSESGEDL
LSPDDTSTGLEEVEQNLHNSFPSSRGRTPGKPMEDTM"
```

3'UTR

repeat_region

BASE COUNT

ORIGIN

11333..>13887

11789..11828

4455 a 2906 c 3164 g 3362 t

4555 a 2906 c 3164 g 3362 t

Query Match

Best Local Similarity

Matches 1090;

51.5%;

91.7%;

Conservative 0;

Score 1030.6;

Pred. No. 1.2e-247;

Mismatches 99;

DB 4;

Indels 0;

Gaps 0;

Length 13887;

872

8385

813

8326

873

GGTACCTACTCATAGATTACTGCAACAGTTCCTCCCTTGAGCTGCAAGAGTTCTTGGCTG

GGAGAAACCACAGATTACTGCAACAGTTCCTCCCTTGAGCTGCAAGAGTTCTTGGCTG

GCTTACAGAGCTGAAACAACTGCCAATGTCTCAGGATGCTACCCGCTAAGGAAAGCT

932

8386

933

GCTTACAGAGCTGAAACAACTGCCAATGTCTCAGGATGCTACCCGCTAAGGAAAGCT

GCTTACAGAGCTGAAACAACTGCCAATGTCTCAGGATGCTACCCGCTAAGGAAAGCT

CCTAGAAGACTCCAGGAGGATAAAGAGCTGATGAAACANTGGCAGACCTCCAGGTGA

992

8445

993

GCTTACAGAGCTGAAACAACTGCCAATGTCTCAGGATGCTACCCGCTAAGGAAAGCT

TCTAGAAGACTCCAGGAGGATAAAGAGCTGATGAAACANTGGCAGACCTCCAGGTGA

AATTGAGCTCAGCAGATGTTTATACAACTGATGATGAAACAACTGGCAGACCTCCAGGAGA

1052

8506

993

AATTGAGCTCAGCAGATGTTTATACAACTGATGATGAAACAACTGGCAGACCTCCAGGAGA

AATCGAAGCTCAGCAGATATCTATCACAACCTGACGAAATGSCCAAAAGTCTCTGAG

8565

1053 ATCCCTGGAGGTTCCGATGCGAGTCTGTGTTACAAAGAGCGTTGGTATACATGAACCTT 1112
1112
8566 ATCCCTGGAGGTTCTGAGATGCGAGCCTGTGTGCAAAGAGCGTTTGGATATACATGAACCTT 8625
1113 CAAAGTGAGTGAACCTTCGGAAAGAGTCTCTCAACATTAGTCCCAATTTGGAGCCAGGTTTC 1172
8626 CAAAGTGAGTGAACCTTCGGAAAGAGTCTCTCAACATTAGTCTCACTTTGGAGCCAGGTTTC 8685
1173 TGACAGTGGAGCGTCTGACACTTCTCTGAGGAACTTCTGTGTGTGGCTACAGCTGAA 1232
8686 TGACAGTGGAGCGTCTGACACTTCTCTGAGGAACTTCTGTGTGTGGCTACAGCTGAA 8745
1233 AGATGATGAATTAAGCGCGCAGCAGCTATTTGAGGCGCACTTTCCAGCAGTTTCAAGACA 1292
8746 AGATGATGAATTAAGCGCGCAGCAGCTATTTGAGGCGCACTTTCCAGCAGTTTCAAGACA 8805
1293 GAACGATGATACATAGGCGCTTCAAGAGGAATTTGAACCTAAGAACCTGTATCATGAG 1352
8806 GAATGATGATACATAGGCGCTTCAAGAGGAATTTGAACCTAAGAACCTGTATCATGAG 8865
1353 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAGGAGTACTAGAGAACT 1412
8866 TACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAGGAGTACTAGAGAACT 8925
1413 CTACAGGAGCCAGAGAGTCTCTCTGAGGAGAGCCCGCAAGATGTCACTCGGCTTCT 1472
8926 CTACAGGAGCCAGAGAGTCTCTCTGAGGAGAGCCCGCAAGATGTCACTCGGCTTCT 8985
1473 ACAGAGCAGGCTGAGAGTCTCAATCTGAGTGGGAAATTTGAACCTCACTTCGGCTGA 1532
8986 ACAGAGCAGGCTGAGAGTCTCAATCTGAGTGGGAAATTTGAACCTCACTTCGGCTGA 9045
1533 CTGCGAGAGAAATAGATGAGACCTTTGAAAGACTTCCAGGAATTTCAAGAGCCACGGA 1592
9046 CTGCGAGAGAAATAGATGAGACCTTTGAAAGACTTCCAGGAATTTCAAGAGCCACGGA 9105
1593 TGAGTGGAGCTCAAGTCTGAGGAGTGTGATCAAGGATCTTGGAGCCGCTGGG 1652
9106 TGAGTGGAGCTCAAGTCTGAGGAGTGTGATCAAGGATCTTGGAGCCGCTGGG 9165
1653 CGATCTCTCTCAATGACTCTCTCAAGATCACTCGAGAACTTCAAGGAGCTTCAAGGAG 9225
9166 TGAGTGGAGCTCAAGTCTGAGGAGTGTGATCAAGGATCTTGGAGCCGCTGGG 9285
1713 AATTGCGCTCTGAAAGAGAGTGTGAGGAGTGTGATCAAGGATCTTGGAGCCGCTGGG 1772
9226 AATTGAGCTCTGAAAGAGAGTGTGAGGAGTGTGATCAAGGATCTTGGAGCCGCTGGG 9285
1773 TTTGGGCAATTCAGCTCTCAAGATCACTCGAGAACTTCAAGGAGCTTCAAGGAG 1832
9286 GTTGGGCAATTCAGCTCTCAAGATCACTCGAGAACTTCAAGGAGCTTCAAGGAG 9345
1833 GAAGCTTCTGAGTGGCGCTTCAAGGAGTGTGATCAAGGATCTTGGAGCCGCTGGG 1892
9346 GAAGCTTCTGAGTGGCGCTTCAAGGAGTGTGATCAAGGATCTTGGAGCCGCTGGG 9405
1893 CTTTGGTCCAGCTCTGAGGAGTGTGATCAAGGATCTTGGAGCCGCTGGG 1952
9406 CTTTGGTCCAGCTCTGAGGAGTGTGATCAAGGATCTTGGAGCCGCTGGG 9465
1953 CATCTCGCCAAACAAAGTGGCTTACTATATCAACACGAGACTCAACA 2001
9466 CATCTCGCCAAACAAAGTGGCTTACTATATCAACACGAGACCCCAACA 9514

RESULT 14
AX306153
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

AX306153
Sequence 904 from Patent WO0188188.
AX306153
AX306153.1 GI:17645441
Mus musculus (house mouse)

ORGANISM Mus musculus
REFERENCE
AUTHORS Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 904 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
1. .13815
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 4577 a 2719 c 3025 g 3494 t
ORIGIN
Query Match 50.2%; Score 1004; DB 6; Length 13815;
Best Local Similarity 90.3%; Pred. No. 5.8e-241;
Matches 1073; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 813 GGTACCTACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGCTG 872
Db 8251 GGAAGAAACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGCTG 8310
QY 873 GCTTACAGAGAGCTGAAACAACTGCCAATGTCTTACAGAGTGTCTACCCGTAAGAAAAGCT 932
Db 8311 GATTACGAGAGAGAGAACTGCCAATGTCTTACAGAGTGTCTTCCGTAAGAGAAGCT 8370
QY 933 CCTAGAGAGCTCCAGGGAGTAAAGAGCTGTGTAACAATGCAAGACCTCCCAAGTGA 992
Db 8371 CCTAGAGAGCTCCAGGGAGTCAAGAGAGCTGTAACAATGCAAGACCTCCCAAGTGA 8430
QY 993 AATTGAGAGCTCACACAGATGTTTATCAACACCTGGATGAAAACAGCCAAAATTCCTGAG 1052
Db 8431 AATTGAGAGCTCACACAGATGTTTATCAACACCTGGATGAAAACAGCCAAAATTCCTGAG 8490
QY 1053 ATCCCTGGAAGGTTCCGATGATGAGTCTCTGTTTACAAGAGCTTTGGATAACATGACTT 1112
Db 8491 ATCCCTGGAAGGTTCCGATGATGAGTCTCTGTTTACAAGAGCTTTGGATAACATGACTT 8550
QY 1113 CAAGTGGAGTGAACCTTCGAAAGAGTCTCTCAACATTTAGGTCCCATTTGGAAGCCAGTTC 1172
Db 8551 CAAGTGGAGTGAACCTTCGAAAGAGTCTCTCAACATTTAGGTCCCATTTGGAAGCCAGTTC 8610
QY 1173 TGACAGTGGAGGCTCTGCACTTTCTGCAAGAACTTCTGTGTGGCTTACAGCTGAA 1232
Db 8611 TGACAGTGGAGGCTCTGCACTTTCTGCAAGAACTTCTGTGTGGCTTACAGCTGAA 8670
QY 1233 AGATGATGAATTAAGCCGCGCAGGACCTTATTTGGAGGCGACTTCCAGCAGTTCAAGAGCA 1292
Db 8671 AGATGATGAATTAAGCCGCGCAGGACCTTATTTGGAGGCGACTTCCAGCAGTTCAAGAGCA 8730
QY 1293 GAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACCTAAGAACTCTGTAATCATGAG 1352
Db 8731 GAATGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACCTAAGAACTCTGTAATCATGAG 8790
QY 1353 TACTCTGAGACTGTACGAATTTCTGACAGAGCAGCTTTGGAGAGCTTAGAGAACT 1412
Db 8791 TACTCTGAGACTGTACGAATTTCTGACAGAGCAGCTTTGGAGAGCTTAGAGAACT 8850
QY 1413 CTACAGGAGCCAGAGAGTGTCTTCTGAGGAGAGCCAGAGATGTCACTCGGCTTCT 1472
Db 8851 CTACAGGAGCCAGAGAGTGTCTTCTGAGGAGAGCCAGAGATGTCACTCGGCTTCT 8910
QY 1473 ACGAAAGCAGGCTGAGGAGGTCATATCTGAGTGGGAAAATTTGAACCTTGGCTCGCTGAGTGA 1532
Db 8911 ACGAAAGCAGGCTGAGGAGGTCATATCTGAGTGGGAAAATTTGAACCTTGGCTCGCTGAGTGA 8970
QY 1533 CTGGCAGAGAAAATAGATGAGAGCCCTTGAAGAGCTTCCAGGAACTTCAAGAGCCACGGA 1592
Db 8971 TTGGCAGAGAAAATAGATGAGAGCTTCTTGAAGAGCTTCCAGGAACTTCAAGAGCTGCGGA 9030
QY 1593 TGAGCTGGAGCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGATCTCTGGCAGCCCGTGGG 1652

Db	9031	TGAACTGGACCTCAAGTTCGGCCCAAGCTGAGTGATCAAGGGATCCTGGCAGCCAGTGGG	9090
Qy	1653	CGATCTCCTCATTGACTCTCTCCAGATACACCTCAGAAAGTCAAGGACATTCGAGGAGA	1712
Db	9091	GGATCTCCTCATTGACTCTCTCCAGATACACCTTGAAAAGTCAAGGACATTCGGGGAGA	9150
Qy	1713	AATTGGCCCTCTGAAGAGAACGTCGAGCACGTCGAATGACCTTGTCTCCGACGCTTACCAC	1772
Db	9151	AATTGCACCTCTAAAGAGAATGTCAATCGTCAATGACCTTGACATCGACGTCGACAC	9210
Qy	1773	TTTGGCATTGACCTCTACCGCTATACCTCAGCACTCTGGAAGACCTGAAACACAGATG	1832
Db	9211	ACTGGGCACTCAGCTCTCACCCTTATACCTCAGCACTCTGGAAGATCTGATACAGATG	9270
Qy	1833	GAAGCTTCTGCAAGTGGCCGTCGAGGACGATGAGGAGCTGATGATGAAGCCACAGGGA	1892
Db	9271	GAGGCTTCTACAGTGGCTGTGGAGACCGTGTCAAGACAGCTGCAATGAAGCCACAGGGA	9330
Qy	1893	CTTTGGTCAGCATCTCAGCACTTTCTTTCCAGCTCTGCAAGGCTCCCTGGGAGAGAGC	1952
Db	9331	CTTTGGTCAGCATCTCAGCACTTTCTTTCCAGCTCTGCAAGGCTCCCTGGGAGAGAGC	9390
Qy	1953	CATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGCTCAAAAC	2000
Db	9391	CATCTCAACCAACAAAGTGCCTTACTATATCAACACGAGACCCCAAC	9438
RESULT 15			
AX538582			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1073; Conservative			
Qy	813	GATACCTACTCATGATTACTGCAAGTTCCTCCCTGCACTGCAAGTTCCTGCTG	872
Db	8251	GGAGAAACTCATGATTACTGCAAGTTCCTCCCTGCACTGCAAGTTCCTGCTG	8310
Qy	873	GCTTACAGAAGCTGAACAACTGCAATGTCTTACAGATGCTTACCCGTAAGGAAGGCT	932
Db	8311	GATTACGGAAGCAGAAACAACTGCCATGTCTTACAGGACGCTTCCCGTAAGGAAGCT	8370
Qy	933	CCTAGAAGACTCCAGGAGTAAAGAGCTGATGAACATGCGCAAGCTCCAGGTGA	992
Db	8371	CCTAGAAGACTCCAGGAGTCAAGAGCTGATGAACATGCGCAAGCTCCAGGAGA	8430
Qy	993	ATTGAGCTCACAGATGTTTATCAACCTGTGATGAACACAGCCAAATTCCTGAG	1052
Db	8431	AATTGAACTCACAGATATCTATCAACCTGTGATGAACATGCGCAAGCTCCAG	8490
Qy	1053	ATCCCTGGGAAGTTCGGATGTCAGTCTCTGTTACAAAGACGTTTGGATACATGAAC	1112

Db	8491	ATCCCTGGAGGTTGGATGAAGCACCCCTGTTACAAAGACGTTTGGATAACATGAATTT	8550
Qy	1113	CAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACATTAGTCCCATTTGGAAGCAGTTTC	1172
Db	8551	CAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACATTAGTCCCATTTGGAAGCAGTTTC	8610
Qy	1173	TGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGGTGTGGCTACAGTGA	1232
Db	8611	TGACCAAGTGAAGCGTTCGATCTTTCTCTCAGGAACCTTCTTGTGTGGCTACAGTGA	8670
Qy	1233	AGATGATGAATTAAGCCGCGAGGACCTTATTCGAGCGGACCTTTCCAGCAGTTCAGAGCA	1292
Db	8671	AGATGATGAATTAAGCCGCGAGGACCTTATTCGAGCGGACCTTTCCAGCAGTTCAGAGCA	8730
Qy	1293	GAACGATGTACATAGAGGCTTTCAAGAGGGAATTTGAATAAAGAACTGTATCATGAG	1352
Db	8731	GAATGATATACATAGAGGCTTTCAAGAGGGAATTTGAATAAAGAACTGTATCATGAG	8790
Qy	1353	TACTTTGAGACTGTAGGAATTTCTGACAGAGGACCTTTTGGAAAGGACTAGAGAACT	1412
Db	8791	TACTTTGAGACTGTAGGAATTTCTGACAGAGGACCTTTTGGAAAGGACTAGAGAACT	8850
Qy	1413	CTACCAAGGACCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCTCTCGGCTTCT	1472
Db	8851	CTACCAAGGACCCAGAGAGCTGCTCTCTGAGGAGAGAGCTCAGAAATGTCTCTCGGCTTCT	8910
Qy	1473	ACGAAAGCGGCTGAGAGGTCATTAAGTGGGAAATTTGAATAAAGAACTGTATCATGAG	1532
Db	8911	ACGAAAGCGGCTGAGAGGTCATTAAGTGGGAAATTTGAATAAAGAACTGTATCATGAG	8970
Qy	1533	CTGCGACAGAAATATAGATGAGACCTTTGAAAGACTCCAGGAACTTTCAAGAGCCACGGA	1592
Db	8971	TTGGCAGAGAAATATAGATGAGCTTTGAAAGACTCCAGGAACTTTCAAGAGCCACGGA	9030
Qy	1593	TGAGCTGCACTCAAGTGGCCAGCTGAGTGTATCAAGGATCTCTGGCAGCCCGTGGG	1652
Db	9031	TGAACCTGCACTCAAGTGGCCAGCTGAGTGTATCAAGGATCTCTGGCAGCCCGTGGG	9090
Qy	1653	CGATCTCTCATTTGACCTCTCTCCAGATCACTCTGAGAAAGTCAAGGACCTTCAGAGAGA	1712
Db	9091	GGATCTCTCATTTGACCTCTCTGAGATCACTTTGAAAGAGTCAAGGACCTTCAGAGAGA	9150
Qy	1713	AATTGCGGCTCTGAAAGAGAACGTCGAGCCAGCTCAATGATCTTGTCTGCCAGCTTACCAC	1772
Db	9151	AATTGCGGCTCTTAAAGAGAAATGTCAATCGTGTCAATGATCTTGTCTGCCAGCTTACCAC	9210
Qy	1773	TTTGGGCACTCAGCTCTCAGCTATACCTCAGCACTCTGAGAGACCTGAAACACAGATG	1832
Db	9211	ACTGGGCACTCAGCTCTCAGCTATACCTCAGCACTCTGAGAGACCTGAAACACAGATG	9270
Qy	1833	GAAGCTTCTGCAAGTGGCCGTCGAGGACCGAGTCAAGGACCTGATGAAGCCACAGGGA	1892
Db	9271	GAGGCTTCTACAGTGGGCTGTGGAGGACCGTGTGAGAGCTGATGAAGCCACAGGGA	9330
Qy	1893	CTTTGGTCCAGATCTCAGCACTTTCTTCCACCTGTCTCCAGGCTCCCTGGGAGAGAGC	1952
Db	9331	CTTTGGTCCAGATCTCAGCACTTTCTTCCACCTGTCTCCAGGCTCCCTGGGAGAGAGC	9390
Qy	1953	CATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGACTCAAAAC	2000
Db	9391	CATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGACTCAAAAC	9438

Search completed: February 1, 2004, 23:50:57
Job time : 7131.37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

M nucleic - nucleic search, using sw model
Run on: February 1, 2004, 11:04:35 ; Search time 3939.6 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000

Perfect score: 2101

Sequence: 1 gagctatgcctacacacagg.....acctcagcactctggaagac 2101

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_estl.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_esthum.*

18: em_estinv.*

19: em_estpln.*

20: em_estvrt.*

21: em_estfun.*

22: em_estmam.*

23: em_estmus.*

24: em_estpro.*

25: em_estrod.*

26: em_estphg.*

27: em_estvrl.*

28: gb_gsel.*

29: gb_gsel.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result	No.	Score	Match	Length	ID	Description
1	1103.4	52.5	3870	11	BC036103	BC036103 Homo sapi
2	866.6	41.2	3056	11	AK044536	AK044536 Mus muscu
3	597.2	28.4	728	14	CB228986	CB228986 AGENCOURT
4	592.4	28.2	834	12	B1729851	B1729851 603349511

5	514.6	24.5	1490	11	BC009242	BC009242 Homo sapi
6	509	24.2	595	14	CB177816	CB177816 is21c01.x
7	491.6	23.4	750	12	B1730168	B1730168 603349711
8	486.6	23.2	579	9	AL121550	AL121550 DKEZ07621
9	424.2	20.2	663	12	BM488464	BM488464 pgmzn.pk0
10	389	18.5	644	13	B0133510	B0133510 603340290
11	343	16.3	2334	11	BC011062	BC011062 Mus muscu
12	324.6	15.4	402	14	CB547284	CB547284 AMGNNUC:S
13	265.8	12.7	772	12	B1250598	B1250598 602993659
14	226	10.8	250	10	BF963618	BF963618 QV2-NN004
15	217.2	10.3	851	13	BU201022	BU201022 603952191
16	212.2	10.1	3051	11	BC036095	BC036095 Homo sapi
17	203.6	9.7	1541	11	AK034383	AK034383 Mus muscu
18	203.6	9.7	3753	11	AK081426	AK081426 Mus muscu
19	192	9.1	778	14	CB524596	CB524596 UI-M-FY0
20	183	8.7	763	14	CB518960	CB518960 UI-M-GH0
21	182.4	8.7	645	9	AW467977	AW467977 he29q08.x
22	177	8.4	646	12	B1289102	B1289102 UI-R-DK0
23	172.8	8.2	835	12	B1553820	B1553820 603190772
24	166.2	7.9	907	13	BU122401	BU122401 603003073
25	165.2	7.9	600	12	B1988528	B1988528 4012-24 M
26	162	7.7	427	10	BB817979	BB817979 CM2-BN027
27	158.2	7.5	615	14	CA377239	CA377239 655768 NC
28	153.2	7.3	885	11	AK020881	AK020881 Mus muscu
29	150.6	7.2	515	28	AZ780914	AZ780914 2M0018010
30	144.8	6.9	434	13	BQ375536	BQ375536 PML-TN012
31	144.6	6.9	675	14	BY706879	BY706879 BY706879
32	142.2	6.8	426	13	BQ319056	BQ319056 IL5-CT051
33	142.2	6.8	481	10	B8557463	B8557463 f103h06.Y
34	141.6	6.7	794	28	BZ097143	BZ097143 CH230-236
35	141	6.7	761	28	B2129893	B2129893 CH230-452
36	140.6	6.7	649	13	BQ209933	BQ209933 UI-R-E90
37	140	6.7	515	10	B8651473	B8651473 BB651473
38	137	6.5	466	14	CD549993	CD549993 B0305E01
39	134	6.4	828	28	BZ126085	BZ126085 CH230-452
40	133.2	6.3	521	28	AZ226551	AZ226551 RPCI-23-9
41	131.6	6.3	410	13	BQ349936	BQ349936 PML-HT034
42	131.6	6.3	784	10	BG212445	BG212445 RST32032
43	129.2	6.1	545	13	EX473852	EX473852 DKFZp686J
44	125.6	6.0	504	12	B1289382	B1289382 UI-R-DK0
45	124.4	5.9	681	13	BU301653	BU301653 603609005

ALIGNMENTS

RESULT 1
BC036103
LOCUS BC036103 3870 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:5274415, mRNA.
ACCESSION BC036103
VERSION BC036103.1 GI:23271310
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3870)
AUTHORS Strausberg R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NTH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-f@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 48 Row: f Column: 7
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: retained intron.

FEATURES

source

1. 3870
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5274415"
 /tissue type="Brain, hippocampus"
 /clone lib="NTH MGC_95"
 /lab host="DH10B"
 /note="vector: pBluescript"

BASE COUNT

1321 a 677 c 763 g 1109 t

ORIGIN

Query Match 52.5%; Score 1103.4; DB 11; Length 3870;
 Best Local Similarity 99.9%; Pred. No. 1.2e-209;
 Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTGACCACTCTGACCTACACGGAGCCCAT 60
 DB 1043 GAGCTATGCTTACACACAGGCTGCTTATGTGACCACTCTGACCTACACGGAGCCCAT 1102

QY 61 TCCTTCACAGCATTTGGAAGCTCTGAAGACAGTTCATTTGGCAGTTCATTTGATGGAGAG 120
 DB 1103 TCCTTCACAGCATTTGGAAGCTCTGAAGACAGTTCATTTGGCAGTTCATTTGATGGAGAG 1162

QY 121 TGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAGTATTTATCGTGGCTTCTTTC 180
 DB 1163 TGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAGTATTTATCGTGGCTTCTTTC 1222

QY 181 TGCTCAGACACATTTGAAGACAGGAGAGATTTCTAATCATGTGGAAGTGGTGAAGA 240
 DB 1223 TGCTCAGACACATTTGAAGACAGGAGAGATTTCTAATCATGTGGAAGTGGTGAAGA 1282

QY 241 CCAGTTTCATCTATGAGGGGTACATGATGATTTGACGCCCATCAGGCCCGGGTTGG 300
 DB 1283 CCAGTTTCATCTATGAGGGGTACATGATGATTTGACGCCCATCAGGCCCGGGTTGG 1342

QY 301 TAAATTTCTAATTTGGGAAGTAACTGATTTGGAACAGGAAATATCAGAGATGAAGA 360
 DB 1343 TAAATTTCTAATTTGGGAAGTAACTGATTTGGAACAGGAAATATCAGAGATGAAGA 1402

QY 361 AACTGAAGTACAAAGCAGATGAATCTCTAAATCAAGATGGGAATGCCCTCAGGGTAGC 420
 DB 1403 AACTGAAGTACAAAGCAGATGAATCTCTAAATCAAGATGGGAATGCCCTCAGGGTAGC 1462

QY 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAAGAACT 480
 DB 1463 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAAGAACT 1522

QY 481 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAAATGGAGGAAGA 540
 DB 1523 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAAATGGAGGAAGA 1582

QY 541 GCCTCTTGACCTGATCTTTGAAGACTTAAAGCGCAAGTACAAACATAGAGTCTTCA 600
 DB 1583 GCCTCTTGACCTGATCTTTGAAGACTTAAAGCGCAAGTACAAACATAGAGTCTTCA 1642

QY 601 AGAGAGTCTAGACAGACAGTCAAGGTCAATTTCTCTCACTCATGTGGTGGTAGT 660
 DB 1643 AGAGAGTCTAGACAGACAGTCAAGGTCAATTTCTCTCACTCATGTGGTGGTAGT 1702

QY 661 TGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTTGGAGAAACAACATTAAGGTATTGGG 720
 DB 1703 TGATGAATCTAGTGGAGATCAAGCAACTGCTGCTTTTGGAGAAACAACATTAAGGTATTGGG 1762

QY 721 AGATCGATGGGCAACATCTCTGTAGATGACAGAGACCGCTGGGTCTTTTACAAGACAT 780
 DB 1763 AGATCGATGGGCAACATCTCTGTAGATGACAGAGACCGCTGGGTCTTTTACAAGACAT 1822

QY 781 CTTTCTCAATTTGGCAACCTCTTACTTGAGAACAGTGCCTTTTGTAGTCATGGCTTTTACA 840
 DB 1823 CTTTCTCAATTTGGCAACCTCTTACTTGAGAACAGTGCCTTTTGTAGTCATGGCTTTTACA 1882

QY 841 AAAAGAAGATGCAAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT 900
 DB 1883 AAAAGAAGATGCAAGTGAACAAGATTCACACAACCTGGCTTTTAAAGATCAAAATGAATGTT 1942

QY 901 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
 DB 1943 ATCAAGTCTTCAAAAACCTGGCGGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2002

QY 961 GGGCAAACTGTATTTCACCTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGC 1020
 DB 2003 GGGCAAACTGTATTTCACCTCAAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGC 2062

QY 1021 CCAGAAGACGGAAGCATGGCTGGATACTTTGCCGGTGGTGGGATAATTTAGTCCAAA 1080
 DB 2063 CCAGAAGACGGAAGCATGGCTGGATACTTTGCCGGTGGTGGGATAATTTAGTCCAAA 2122

QY 1081 ACTTGAAAAGAGTACAGCACAGACT 1105
 DB 2123 ACTTGAAAAGAGTACAGCACAGACT 2147

RESULT 2
 AK044536 3056 bp mRNA linear HTC 05-DEC-2002
 LOCUS Mus musculus adult retina cDNA, RIKEN full-length enriched library,
 DEFINITION clone:A930019F21 product:dystrophin, muscular dystrophy, full insert sequence.
 ACCESSION AK044536
 VERSION AK044536.1 GI:26090404
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE

4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carinci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C., Fleischer, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Rongwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.

5
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6921), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851

REFERENCE

5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 AUTHORS 6 (bases 1 to 3056)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hoshizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Oheato, N., Okazaki, Y., Saito, R., Saich, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Retina RNA was provided by Dr. Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers

FEATURES

1. .3056
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:A930019P21"
 /db_xref="taxon:10090"
 /clone="A930019P21"
 /tissue type="retina"
 /clone lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"

misc_feature 1. .3056
 /note="dystrophin, muscular dystrophy (MGD|MGI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)"
 BASE COUNT 1072 a 557 c 810 t
 ORIGIN
 Query Match 41.2%; Score 866.6; DB 11; Length 3056;
 Best Local Similarity 86.5%; Pred. No. 1.9e-162;
 Matches 956; Conservative 0; Mismatches 149; Indels 0; Caps 0;
 Qy 1 GAGCTATGCTACACACAGCGTCTTATGTCACCACTCTGACCCCTACAGGAGCCCAATT 60
 Db GAGTTATGCTTACACAGCGTCTTATGTCACCACTCTGATTCACACAGAGCCCAATT 1003
 Qy 61 TCCTTACACAGCATTTGGAAGCTCTCTGAAGACAGTCAATTTGGCAGTTCAATTGATGAGAG 120
 Db TCCTTACACAGCATTTGGAAGCTCTCTGAAGACAGTCAATTTGATGATGATGATGATGATG 1063
 Qy 121 TGAAGTAAACCTGACCGGTATCAACACAGCTTTAGAGAGAGTATTTCTGTTGGTCTTTTC 180
 Db GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAGAGAGTATTTCTGTTGGTCTTTTC 1123
 Qy 181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA 240
 Db TGCGAGGATACATTTGGAGCACAAGGAGATTTCAATGATGTGGAAGTGTGGAAGA 1183
 Qy 241 CGAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCGGGTTGG 300
 Db ACAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCGGGTTGG 1243
 Qy 301 TAATATTTACAAATTTGGGAAGTAAAGCTGATTTGGAACAGGAAAATTTATCAGAAGTGAAGA 360
 Db TAATATTTACAAATTTGGGAAGTAAAGCTGATTTGGAACAGGAAAATTTATCAGAAGTGAAGA 1303
 Qy 361 AACTGAAGTACAAAGACAGATGAATCTCTAAATCAAGATGGAATGCTCAGGGTAGC 420
 Db AGCTGAAGTCAAAAGACAAATGAATCTCTAAATCAAGATGGAATGCTCAGGGTAGC 1363
 Qy 421 TAGCTGGAAGTCAAAAGACAAATGAATCTCTAAATCAAGATGGAATGCTCAGGGTAGC 480
 Db TAGCTGGAAGTCAAAAGACAAATGAATCTCTAAATCAAGATGGAATGCTCAGGGTAGC 1423
 Qy 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACAAAGAAAATGGAGGAAGA 540
 Db AAAGAGTGAATGACTGGCTAAACAAAACAGAGAAAGAACAAAGAAAATGGAGGAAGA 1483
 Qy 541 GCCTTTGAGACCTGATCTTGAAGACCTTAAAGCCGAGTACAAACATAGAGTGGTTC 600
 Db GCCTTTGAGACCTGATCTTGAAGACCTTAAAGCCGAGTACAAACATAGAGTGGTTC 1543
 Qy 601 AGAAGATCTAGAACAAAGACAAAGTCAAGGTCAATTTCTCACTCACTCACTGGTGGT 660
 Db AGAAGATCTAGAACAAAGACAAAGTCAAGGTCAATTTCTCACTCACTCACTGGTGGT 1603
 Qy 661 TGATGAATCTAGTGGAGATCAGCCAACTGCTGTTTGAAGAACCACTTAAGGTATGGG 720
 Db TGATGAATCTAGTGGAGATCAGCCAACTGCTGTTTGAAGAACCACTTAAGGTATGGG 1663
 Qy 721 AGATCGATGGGAAAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTCAAGACAT 780
 Db AGATCGATGGGAAAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTCAAGACAT 1723
 Qy 781 CCTTCTCAATGGCAACGCTTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTCAGA 840
 Db CCTTCTCAATGGCAACGCTTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTCAGA 1783
 Qy 841 AAAGAGAGTGAAGTGAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAATGTT 900
 Db AAAGAGAGTGAAGTGAAGATTCACACAACTGGCTTTTAAAGATCAAAATGAATGTT 1843
 Qy 901 ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
 Db ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1903

```

QY 961 GGGCAAACTGTATTCACTCAAAAGAGTCTTCTTCAACACTGAAGATAAGTCAGTGAC 1020
DB 1904 GGAATAAATTAAGTTCATCAATCAAGATCTACTTTCGGCACTGAAAATTAAGTCAGTGAC 1963
QY 1021 CCAGAAAGCGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAAATTAGTCCAAAA 1080
DB 1964 TCAAAAGATGGAATCTCGATGGAAAACCTTTGCACACGTTGGGACAAATTAAACCCAAAA 2023
QY 1081 ACTTGAAGAGTACAGCACAGACT 1105
DB 2024 ACTTGAAGAGTTCAGCACAAATT 2048

CB228986 728 bp mRNA linear EST 10-FEB-2003
AGENCOURT_11499247 NICHD_Rh_Ovi Macaca mulatta cDNA clone
IMAGE:6884820 5', mRNA sequence.
CB228986
CB228986.1 GI:28280564
EST.
Macaca mulatta (rhesus monkey)
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
1 (bases 1 to 728)
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
cDNA Library Preparation: CLONTECH
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3135 row: e column: 11
High quality sequence stop: 583.
Location/Qualifiers
1..728
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_Rh_Ovi"
/note="Organ: ovary; Vector: pDNR-LIB; Site_1: Sfi I;
Site_2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal and menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."
BASE COUNT 256 a 127 c 162 g 182 t 1 others
ORIGIN
Query Match 28.4%; Score 597.2; DB 14; Length 728;
Best Local Similarity 97.0%; Pred. No. 8.6e-109;
Matches 619; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 295 GGTGGTAATATTCTACAAATGGGAAGTAAGCTGATTGGAACAGGAAATATTCAGAGA 354
DB 1 GGTGGTAATATTCTACAAATGGGAAGTCAGCTGATTGGAACAGGAAATATTCAGAGA 60
QY 355 TGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGATGGGAATGCTCAG 414
DB 61 TGAAGAACTGAAGTACAAGACAGATGAATCTCTAAATTCAAGATGGGAATGCTCAG 120
QY 415 GGTAGCTACATGGGAACAAACAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCA 474

```

```

DB 121 GGTAGCTACATGGGAACAAACAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCA 180
QY 475 GAAACTGAAGAGTTGAATGACTGCTTAACAAACAGCAAGAAAGCAACAGGAAATGGA 534
DB 181 GAAACTGAAGAGTTGAATGACTGCTGCTACAAACAGCAAGAAAGCAACAGGAAATGGA 240
QY 535 GGAAGAGCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACATAAGGT 594
DB 241 GAAAGAACCCCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACATAAGGT 300
QY 595 GCTTCAAGAGATCTAGAACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 654
DB 301 GCTTCAAGAGATCTAGAACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 360
QY 655 GGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTCTTTTGAAGAACCAACTTAAGGT 714
DB 361 GGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTCTTTTGAAGAACCAACTTAAGGT 420
QY 715 ATGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTCTCTTTTACA 774
DB 421 ATGGGAGATCGATGGGCAACATCTGTAGATGGACAGAACCGCTGGTCTCTTTTACA 480
QY 775 AGACATCTCTTCAATGCGCAAGTCTTACTGAAGACAGTGCCTTTTGTGATGCT 834
DB 481 AGACATCTCTTCAATGCGCAAGTCTTACTGAAGACAGTGCCTTTTGTGATGCT 540
QY 835 TTCAAGAAAGAGATGCGATGCAACAGATTCACAACT-GGCTTTAAAGATCAAAATG 893
DB 541 TTCAAGAAAGAGATGCGTGAACAGGATTCACAACTGGCTTTAAGGATCAAAATG 600
QY 894 AAATGTTATCAAGTCTTCAAAACCTGGCGCTTTTAAA 931
DB 601 AAATGTTATCAAGTCTTCAAAACCTGGCGCTTATGTA 638

RESULT 4
BI729851 834 bp mRNA linear EST 20-SEP-2001
LOCUS 603349511F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357162 5',
DEFINITION mRNA sequence.
BI729851
BI729851.1 GI:15706864
EST.
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 834)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1908 row: e column: 03
High quality sequence stop: 796.
Location/Qualifiers
1..834
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357162"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;

```

Site 2: Salt; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."

```

BASE COUNT      275 a 173 c 199 g 187 t
ORIGIN
Query Match      28.2%; Score 592.4; DB 12; Length 834;
Best Local Similarity 88.4%; Pred. No. 7.9e-108;
Matches 677; Conservative 0; Mismatches 86; Indels 3; Gaps 3;
2y 1065 ATAATTAGTCCAAACCTTTGAAAGAGTACAGCAGACTCATAGATTATGCAACAGT 1124
   |||
3b 44 ATAAAGAGTAAGTACGACAGAGGCTGCTTGGAAAGAACTCATAGATTACTGCGCAGT 103
   |||
2y 1125 TCCTCCCTGACCTGGAAAGTTCTTCCTGCTGCTTACAGAGCTCAACAACTGCCAATG 1184
   |||
3b 104 TCCTCTGACCTGGAGAGTTCTTCCTGCTGCTTACAGAGCTCAACAACTGCCAATG 163
   |||
2y 1185 TCCTACAGGATGCTACCCGTAAGGAAAGCTCTCTAGAGACTCCCAAGGGAGTAAAGAGC 1244
   |||
3b 164 TCCTACAGGAGCTTCCCGTAAGGAGAGCTCTCTAGAGACTCC-AGGAGTCAAGAGC 222
   |||
2y 1245 TGAATGAACAAATGGCAGACCTCCAGGTGAATTCAGCTCACACAGATGTTATCA 1304
   |||
3b 223 TGAATGAACAAATGGCAGACCTCCAGGTGAATTCAGCTCACACAGATGTTATCA 282
   |||
2y 1305 ACCTGATGAAACACGCCAAAATCCTCAGATCCCTGGAAGGTTCCAGATGATGAGTCC 1364
   |||
3b 283 AATCTGATGAAATGSCCAAAATCCTCAGATCCCTGGAAGGTTCCGATGAGCAGCC 342
   |||
2y 1365 TGTACAAAGAGCTTTGGATATACATGAATTCAGTGAAGTGAATTCGGAAGAGTCTC 1424
   |||
3b 343 TGTACAAAGAGCTTTGGATATACATGAATTCAGTGAAGTGAATTCGGAAGAGTCTC 402
   |||
2y 1425 TCAACATTAGTCCCTATTTGGAGCCAGTCTTACAGTGAAGGCTCTGCACCTTTCTC 1484
   |||
3b 403 TCAACATTAGTCCCTATTTGGAGCCAGTCTTACAGTGAAGGCTCTGCACCTTTCTC 462
   |||
2y 1485 TGAGGAACCTTCGTGTGGCTACAGCTGAAGATGATGAATTAAGCGGAGGACCTTA 1544
   |||
3b 463 TTGAGGAACCTTCGTGTGGCTACAGCTGAAGATGATGAATTAAGCGGAGGACCTTA 522
   |||
2y 1545 TTGAGGCGGACTTTCCAGCAGTTTCAGAGCAGAGCTACATAGGCGCTTCAAGAGGG 1604
   |||
3b 523 TCGGTGGTATTTCCAGCAGTTTCAGAGCAGAGCTACATAGGCGCTTCAAGAGGG 582
   |||
2y 1605 AATTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGA 1664
   |||
3b 583 AATTGAAAACCTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTGAGATATTTCTGA 642
   |||
2y 1665 CAGAGCAGCTTTGGAGAGCTTAGAGAACTC-TACCAGAGCCCGAGAGCTGCTCTCT 1723
   |||
3b 643 CAGAGCAGCTTTGGAGAGCTTAGAGAACTCTTACCAGAGCCCGAGAGAACTGCTCTCT 702
   |||
2y 1724 GAGGAGAGAGCCAGAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACT 1783
   |||
3b 703 GAAGAAGAGCTCAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAACGCT 762
   |||
2y 1784 GAGTGGGAAAATTGAACCTGCACTC-CGCTGACTGGCAGAGAAA 1828
   |||
3b 763 GAATGGGACCAATTGAACCTGCGCTCAAGCTGATTGGCAGAGAAA 808
   |||

```

```

RESULT 5
BC009242      1490 bp mRNA linear HTC 04-MAR-2003
LOCUS
DEFINITION
Homo sapiens, dystrophin (muscular dystrophy, Duchenne and Becker
types), includes DXS142, DXS164, DXS206, DXS230, DXS239, DXS268,
DXS269, DXS270, DXS272, clone IMAGE:3029414, mRNA.
BC009242
ACCESSION
VERSION
KEYWORDS
BC009242.1 GI:14714379
HTC.

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1490)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 10 Row: j Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5032284
This clone has the following problem: retained intron.

FEATURES
Location/Qualifiers
1..1490
1. organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="3029414"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC 17"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
BASE COUNT 505 a 299 c 328 g 358 t
ORIGIN
Query Match 24.5%; Score 514.6; DB 11; Length 1490;
Best Local Similarity 99.2%; Pred. No. 2.8e-92;
Matches 517; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GAGTATGCTTACACAGAGCTGCTTATGTACACCTCTGACCTCTACAGGAGCCCAIT 60
Db 969 GAGTATGCTTACACAGAGCTGCTTATGTACACCTCTGACCTCTACAGGAGCCCAIT 1028
Qy 61 TCCTTACACAGCTTTGGAGCTCTCTGAAGACAGTCAATTTGGCAGTTCATTGATGAGAG 120
Db 1029 TCCTTACACAGCTTTGGAGCTCTCTGAAGACAGTCAATTTGGCAGTTCATTGATGAGAG 1088
Qy 121 TGAAGTAAACCTGACCGTTATCAAAACAGCTTTAGAAAGTATTATCGTGGCTTTCTTC 180
Db 1089 TGAAGTAAACCTGACCGTTATCAAAACAGCTTTAGAAAGTATTATCGTGGCTTTCTTC 1148
Qy 181 TGCTGAGGACACATTGCAAGCAAGAGAGATTTCTTAATGATGTGGAGTGTGGAAGA 240
Db 1149 TGCTGAGGACACATTGCAAGCAAGAGAGATTTCTTAATGATGTGGAGTGTGGAAGA 1208
Qy 241 CCAGTTTTCATACCTCATGAGGGGTACATGATGATGATGATGATGATGATGATGATGATG 300
Db 1209 CCAGTTTTCATACCTCATGAGGGGTACATGATGATGATGATGATGATGATGATGATGATG 1268
Qy 301 TAATATTCTCAATTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 360
Db 1269 TAATATTCTCAATTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1328
Qy 361 AACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 420
Db 1329 AACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1388

QY 421 TAGCATGAAAAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGAACTCAGAACT 480
 DB 1389 TAGATGAAAAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGAACTCAGAACT 1448
 QY 481 GAAAGAGTTGAATGACTGGCTGCTAACAAAAACAGAAAGAA 521
 DB 1449 GAAAGAGTTGAATGACTGGCTGCTAACAAAAAAGAAAAA 1489

RESULT 6
 CBI77816/c
 LOCUS
 DEFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3', mRNA sequence.
 similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ;, mRNA sequence.
 ACCESSION CBI77816
 VERSION CBI77816.1 GI:28186206
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 595)
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, L., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, B., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y., and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished
 Other ESTs: is21c01.y1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 448.
 Location/Qualifiers
 1..595
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6553129"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_libs="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1: NotI, Site 2: XhoI; cDNA made by oligo-dT priming. Size selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."
 BASE COUNT 146 a 136 c 116 g 197 t
 ORIGIN

Query Match 24.2%; Score 509; DB 14; Length 595;
 Best Local Similarity 100.0%; Pred. No. 3.3e-91;
 Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTATGCTACACACAGGCTGCTTATGTACACCTCTGACCCCTACAGGACCCCACT 60

DB 509 GAGCTATGCTACACACAGGCTGCTTATGTACACCTCTGACCCCTACAGGACCCCACT 450
 QY 61 TCCTTCACAGCATTTGGAGCTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
 DB 449 TCCTTCACAGCATTTGGAGCTCTCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 390
 QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTGAAGAAGTATTATCGTGGCTTCTTTC 180
 DB 389 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTGAAGAAGTATTATCGTGGCTTCTTTC 330
 QY 181 TCCTGAGGACACATTCGACGACACAGGAGAGATTTCTAATCATGTGGAAGTGGTGAAGA 240
 DB 329 TGCTGAGGACACATTCGACGACACAGGAGAGATTTCTAATCATGTGGAAGTGGTGAAGA 270
 QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGATGATTTGACAGCCCATCAGGCCCGGGTTGG 300
 DB 269 CCAGTTTTCATCTCATGAGGGGTACATGATGATGATTTGACAGCCCATCAGGCCCGGGTTGG 210
 QY 301 TAAATTTCTCAATTTGGGAAGTAACTGATTTGGAACAGGAAAAATTCAGAAAGATGAAGA 360
 DB 209 TAAATTTCTCAATTTGGGAAGTAACTGATTTGGAACAGGAAAAATTCAGAAAGATGAAGA 150
 QY 361 AACTCAAGTACAAAGACGAGATGATCTCTCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 420
 DB 149 AACTCAAGTACAAAGACGAGATGATCTCTCTAAATTCAGATGGGAATGCCCTCAGGGTAGC 90
 QY 421 TAGCATGAAAAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGAACTCAGAACT 480
 DB 89 TAGCATGAAAAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGAACTCAGAACT 30
 QY 481 GAAAGAGTTGAATGACTGGCTAACAAAA 509
 DB 29 GAAAGAGTTGAATGACTGGCTAACAAAA 1

RESULT 7
 CBI730168
 LOCUS
 DEFINITION B1730168 750 bp mRNA linear EST 20-SEP-2001
 mRNA sequence.
 ACCESSION B1730168
 VERSION B1730168.1 GI:15707181
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.mgi.nhl.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Place: LLNMI1908 row: f column: 04
 High quality sequence stop: 747.
 Location/Qualifiers
 1..750
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:5357187"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_94"
 /note="Organ: eye; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed."

FEATURES
 source

Average insert size 3.3 Kb. Library enriched for full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 240 a 161 c 178 g 170 t 1 others

ORIGIN

Query Match 23.4%; Score 491.6; DB 12; Length 750;
Best Local Similarity 89.3%; Pred. No. 9.9e-88;
Matches 598; Conservative 0; Mismatches 64; Indels 8; Gaps 6;
QY 1098 CACGAGCTCATAGTACTGACAGTTCCTCCCTGACCTGGAAAGTTCTTGCTGG- 1156
Db 80 CAGAACTCATAGTACTGACAGTTCCTCTGACCTGGAAAGTTCTTGCTGA 139
QY 1157 CTTACGAACTGAAAC--AACTGCCAATGTCTCAGAGTGTATCCCG-TAAGGAAAG 1213
Db 140 TTTACGGAAGCAAGAAACAGAGTGCCTCAATGTCTCAGAGAGCTTCCCGCTAAGGAGAG 199
QY 1214 CTCTAGAGACTCCCAAGG--AGTAAAGAGTGTATGAAACATGGCAAGACTCCCAAG 1271
Db 200 CTCTAGCAAGACTCCCAAGGAGCTGATGAAACATGGCAAGACTCCCAAG 259
QY 1272 GTGAAA-TTGAAGCTCACAGAGTGTATGAAACCTGATGAAACAGCCAAATATC 1330
Db 260 GAGAACTTGAACTCACAGATATCTATCAAACTTGTGAAATGCCAAATATC 319
QY 1331 CTGAGATCCCTGGAAGTTCGAGTGTAGTGTCTCTTACAAAGAGCTTTGGATACATG 1390
Db 320 CTGAGATCCCTGGAAGTTCGAGTGTAGTGTCTCTTACAAAGAGCTTTGGATACATG 379
QY 1391 AACTCAAGTGGAGTGAACCTTCGGAAGAGTCTCTCAACATAGGTCCTATTTGGAAGCC 1450
Db 380 AATTTCAGTGGAGTGAACCTTCGGAAGAGTCTCTCAACATAGGTCCTATTTGGAAGCA 439
QY 1451 AGTCTGACAGTGGAGGCTCTGCACCTTTCTCTCAGGAACTTCTGCTGCTGCTACAG 1510
Db 440 AGTCTGACAGTGGAGGCTTTGCTCTCTCTCAGGAACTTCTGCTGCTGCTACAG 499
QY 1511 CTGAAGATCATGANTTAAGCCGAGGAGCTATTTGGAGGAGCTTTCCAGAGTTCAG 1570
Db 500 CTGAAGATCATGANTTAAGCCGAGGAGCTATTTGGAGGAGCTTTCCAGAGTTCAG 559
QY 1571 AAGCA-GAAGCATGTACATAGGCTTCAGAGGAGTTCAGGAGGAGTTCAGGAGGAGTTCAG 1629
Db 560 AAGCAAGATGTATACATAGGCTTCAGAGGAGTTCAGGAGGAGTTCAGGAGGAGTTCAG 619
QY 1630 CATGAGTACTCTTGAGCTGTAGCAATATTTCTGACAGAGAGCTTTGGAAGGAGTTCAG 1689
Db 620 CATGAGTACTCTTGAGCTGTAGCAATATTTCTGACAGAGAGCTTTGGAAGGAGTTCAG 679
QY 1690 GAAACTCTACAGAGGAGGAGGAGTTCAGGAGGAGTTCAGGAGGAGTTCAGGAGGAGTTCAG 1749
Db 680 GAAACTCTACAGAGGAGGAGGAGTTCAGGAGGAGTTCAGGAGGAGTTCAGGAGGAGTTCAG 739
QY 1750 GCTTCTACGA 1759
Db 740 GCTTCTACGA 749

RESULT 8
AL121550
LOCUS
DEFINITION
DXF2P762L078 r1 762 (synonym: hm12) Homo sapiens cDNA clone
KEYWORDS
EST.
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 579)
Ottewaelde, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and

Wiemann, S.

EST (Ottewaelde, et al.)

Unpublished

Contact: Ottewaelde B

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by MedGenomix (Martinsried/Germany) within the cDNA

sequencing consortium of the German Genome Project. No s1 sequence

available.

This clone (DKFZ762L078) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1. 579

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZ762L078"

/tissue_type="melanoma (MeWo cell line)"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="762 (synonym: hm12)"

/notes="vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 180 a 128 c 144 g 127 t

ORIGIN

Query Match

23.2%; Score 486.6; DB 9; Length 579;

Best Local Similarity 97.2%; Pred. No. 9.5e-87;

Matches 495; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1260 AAGACCTCCAGGTGAAATTTGAAGCTCACAGATGTTTATCAACCTGGATGAAACA 1319

Db 71 AAGACCTCCAGGTGAAATTTGAAGCTCACAGATGTTTATCAACCTGGATGAAACA 130

QY 1320 GCCAAAAATCTCCAGATCCCTGGAAGTTCCGATGATGAGTCTCTTCAAAAGAGCTT 1379

Db 131 GCCAAAAATCTCCAGATCCCTGGAAGTTCCGATGATGAGTCTCTTCAAAAGAGCTT 190

QY 1380 TGGATAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGTCCC 1439

Db 191 TGGATAACATGAACTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGTCCC 250

QY 1440 ATTTGGAGCCAGTCTCGACCGTGGAGCGTCTGCACCTTTCTCTGCAGAACTCTGG 1499

Db 251 ATTTGGAGCCAGTCTCGACCGTGGAGCGTCTGCACCTTTCTCTGCAGAACTCTGG 310

QY 1500 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACCTATTGGAGGCGACTTC 1559

Db 311 TGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGACCTATTGGAGGCGACTTC 370

QY 1560 CAGCAGTTGAGAGCAAGATGTACATAGGCGCTTCAGAGGAGTTCAGAGGAGTTCAGAG 1619

Db 371 CAGCAGTTGAGAGCAAGATGTACATAGGCGCTTCAGAGGAGTTCAGAGGAGTTCAGAG 430

QY 1620 AACCTGTAATCATGAGTACTCTTGAGACTGTAGAAATTTCTGACAGAGCAGCCTTTGG 1679

Db 431 AACCTGTAATCATGAGTACTCTTGAGACTGTAGAAATTTCTGACAGAGCAGCCTTTGG 490

QY 1680 AAGCAGTACAGAACTCTACAGAGCCCAAGAGAGTGCCTCTCTGAGGAGAGAGCCAGA 1739

Db 491 AAGCAGTACAGAACTCTACAGAGCCCAAGAGAGTGCCTCTCTGAGGAGAGAGCCAGA 550

QY 1740 ATGTCACCTCGGCTTCTACGAAAGCAGGCT 1768

Db 551 AATGTCACCTCGGCTTCTACGAAAGCAGGCT 579

RESULT 9

BM488464

LOCUS

663 bp mRNA linear EST 07-FEB-2002

```

DEFINITION      pgm2n.pk007.117 Normalized Chicken Breast Muscle, Leg Muscle, and
                  Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus CDNA
                  clone pgm2n.pk007.117 5' similar to sp|E11533|DMD_CHICK_Dystrophin
                  p1r|S0241 dystrophin, muscle - chicken emb|CAA31746.1| (X13369)
                  dystrophin (AA 1 - 3660) [Gallus gallus], mRNA sequence.
ACCESSION       BM488464
VERSION         BM488464.1 GI:18609395
KEYWORDS        EST.
SOURCE          Gallus gallus (chicken)
ORGANISM        Gallus gallus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
                Phasianinae; Gallus.
REFERENCE       1 (bases 1 to 663)
AUTHORS         Cogburn, L.A. and Monsonogo-Ornan, E.
TITLE           ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
                Epiphyseal Growth Plate cDNA library, USDA/IFAPS Animal Genome
                Project
JOURNAL         Unpublished
COMMENT         Contact: Larry A. Cogburn
                University of Delaware
                Townsend Hall, Newark, DE 19717, USA
                Tel: 302-831-1335
                Fax: 302-831-2822
                Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES        Location/Qualifiers
                1..663
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /strain="Commercial broiler and Ottawa Res. Centre
                Strains 90 & 21"
                /db_xref="taxon:9031"
                /clone="pgm2n.pk007.117"
                /sex="Male and Female"
                /tissue_type="Breast muscle, leg muscle and epiphyseal
                growth Plate"
                /dev_stage="Breast, leg; Embryo (dl9); post-hatch (dl1,3,5,7,9
                ,11 weeks); growth plate (dl7, dl14 post-hatch)"
                /lab_host="E. coli EMDH10B"
                /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
                and Epiphyseal Growth Plate cDNA library (pgm2n)"
                /note="Vector: pCMVSPORT6; Library made from equivalent
                pools of total RNA isolated from each tissue (embryonic
                muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
                plate 33.3% of the final RNA pool). Single pass sequencing
                from 5'-end"
BASE COUNT      209 a 132 c 172 g 144 t 6 others
ORIGIN
Query Match     20.2%; Score 424.2; DB 12; Length 663;
Best Local Similarity 78.1%; Pred. No. 2.7e-74;
Matches 518; Conservative 0; Mismatches 144; Indels 1; Gaps 1;

QY 1243 GGTGATGAACCAATGCGAAGACTTCAAGGTGAAATTCAGGCTCACAGATGTTTATCA 1302
    |||
Db 1 GCTCATGAGCAGTGCAGGATCTACAGCAGAAATTCAGTGCACATCTGCATCT-TNN 59

QY 1303 CAACTGGATGAACAGCAAAATTCAGATCCCTGGAGTTCGATGTCAGT 1362
    |||
Db 60 CAACTGGATGAACAGGGCAGAAATTCAGATCCCTGGAGGCTCAGAGGATGCTGT 119

QY 1363 CTTGTTACAAAGACGTTTGGATAACATGAATTCAGTGGAGTGAATTCGAAAGAGTC 1422
    |||
Db 120 CTTGTTGAGAGACGTTCTGGATAACATGAATTCAGTGGAGTGAATTCGAAAGAGTC 179

QY 1423 TCTCAACATTAGTCCCATTTGAGACCCAGTCTGACCAAGTGGAGCGTCTGCACCTTTC 1482
    |||
Db 180 TCTAAACATTAGATCTCAATTTGAGACCCAGCAGACAGCAGTGGAGCGTTTACATCTCTC 239

QY 1483 TCTGCGAGGACCTTCTGGTGGCTCAGCTGAAAGATGATGAATTAAGCCGCGAGCACC 1542
    |||
Db 240 TCTTCAGGAACTTTTGGCATGGCTGCAATGAAGGAGGATGAATTAACACAGCAGCACC 299

```

```

QY 1543 TATTGGAGCGCACTTTCCAGCAGTTTCAGAAGCAGAACGATGTACATAGGGCCTTCAAGAG 1602
    |||
Db 300 CATTTGGTGGAGATATTCACACTGTGCAGAGCAGATGATGTTCATAGGACTTTCAGAGAG 359
    |||
QY 1603 GGAATTTGAAAACCTAAAGAACCTGTATCATATGATGACTCTTGTAGACTGTACGATATTTCT 1662
    |||
Db 360 GGAGCTGAAAACAAAAGAACCTGTTTATCATGAATGACTTGTAGACTGTGCGACTTTCTCT 419
    |||
QY 1663 GACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCCTCC 1722
    |||
Db 420 GGAGATCAACACAGTAGAGGAGCTGGAAAGGTCTTATCCAGAACCAAGAGACTTATCACC 479
    |||
QY 1723 TGAGGAGAGAGCCAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGGTCAATATAC 1782
    |||
Db 480 TGAGGAGAGAGCCAGATGTCACCTAAAGTTCTCCGAGGCAAGCAGATGATGTGAGNAC 539
    |||
QY 1783 TGAGTGGGAAATTTGAACCTGCATCTCCGCTCAGCTGGCAGAGAAAATAGATGAGACCT 1842
    |||
Db 540 TGAGTGGGATAAGCTAAATCTAGTTCTGCTGATTGGCAAAGAGATAGATGATGCTCT 599
    |||
QY 1843 TGAAGAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGCAAGC 1902
    |||
Db 600 TGAAGAGACTCGAGGCTTTCAGGAGGCAATGATGAATCTNNCTGAAACTGCGCAGGC 659
    |||
QY 1903 TGA 1905
    |||
Db 660 TGA 662

RESULT 10
BU313510
LOCUS          BU313510.1 GI:25821511
DEFINITION     Gallus gallus (chicken)
ACCESSION      BU313510
VERSION        BU313510.1
KEYWORDS       EST.
SOURCE         Gallus gallus (chicken)
ORGANISM       Gallus gallus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
                Phasianinae; Gallus.
REFERENCE       1 (bases 1 to 644)
AUTHORS         Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
                Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
TITLE           A Comprehensive Collection of Chicken CDNAs
JOURNAL         Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE        2235534
PUBMED         1245392
COMMENT        Contact: Simon Hubbard
                Department of Biomolecular Sciences
                University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES        Location/Qualifiers
                1..644
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /strain="Compton line 15"
                /db_xref="taxon:9031"
                /clone="CHEST508c24"
                /sex="Female"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="CSEQCHN61"
                /note="Organ: heart; Vector: pBluescript II KS(+); Site_1:
                EcoRI; Site_2: NotI; This normalized library was
                constructed from 1 million independent clones. cDNA
                synthesis was initiated using an oligo(dT) primer, using
                methylated C in the first strand synthesis reaction."

```

Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

```
BASE COUNT      192 a  141 c  168 g  142 t  1 others
ORIGIN
Query Match      18.5%; Score 389; DB 13; Length 644;
Best Local Similarity 75.5%; Pred. No. 2.8e-67;
Matches 482; Conservative 0; Mismatches 156; Indels 0; Gaps 0;
NY 1449 CCAGTTCGACAGGGAAGCGTTCGACCTTCTCTCAGCAACTCTCTGTTGGGTAC 1508
b 1449 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
b 7 CCAGCACAGACCGAGTGGAGCGTTTACATCTCTCTCTTCAGAACTTTTGGCATGGTGC 66
Y 1509 AGCTGAAAGATGATGAATTAAGCCGCGAGGACCTATTGGAGCGGACTTTCCAGCAATTC 1568
b 67 AATTGAAGGAGGATCAATTAACACAGCAAGCACCCATTGGTGGAGATATCCCACTGTGC 126
Y 1569 AGAACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1628
b 127 AGAAGCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 186
Y 1629 TCATCAGTACTCTTCGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGACTAG 1688
b 187 TCATGAATGCACTTCGAGACTGTGCGACTCTCTCTGCGAGATCAACCACTGAGGACTGG 246
Y 1689 AGAACTCTACAGAGGCGCAGAGAGTGCCTCTCTGAGAGAGAGCCAGAAATGCTACTC 1748
b 247 AAAAGGTCTATCCAGAACCAAGAGACCTTACCTGAGGAGAGGCCCGAGAAATGTCTACTA 306
Y 1749 GCGTCTACGAAGAGGCTGAGGAGGTCAATPACTGAGTGGGAAATTTGAACCTGCACT 1808
b 307 AAGTCTCCGAAGGACAGAGATGATGTGAGACTGATGAGTGGATTAATCTACGTT 366
Y 1809 CCGCTGACTGGCAGAGAAATATAGATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGG 1868
b 367 CTGCTGATTGGCAAAAGATAGATGATGCTCTTTGAAAGACTGCAGGGTCTTCAGGAGG 426
Y 1869 CCACGATGAGTGCACCTACAGCTGCGCCAGCTGAGTATCAAGGATCTCTGGCAGC 1928
b 427 CAATGATGAATAGACCTGGAACCTGCGCCAGCTGAGAGATTCAGGATTCCTGGCAGC 486
Y 1929 CCGTGGCGGATCTCTCTCACTCTCTCCAGATCACTCGAGAAAGTCAAGCACTTC 1988
b 487 CAGTGGGGATCTGCTGATAGACTCTCTGAGGATCACTTAGAAAAGTCAAGGTTTATC 546
Y 1989 GAGGAGAAATGGGCTCTGAAAGAGAGCTGAGCAGCTCAATCACTTCTGCTGCCAGC 2048
b 547 GAGCAGAAATGGTGGCCCTTANAGAGAGGTCATCAAGTCAATGAGTGGCTCACCGGT 606
Y 2049 TTACCACTTTGGSCATTCAGCTCTCACCGTATAACCTC 2086
b 607 TCGCTCCCTGATATTCAGTCTCTCCCAATACACTCTC 644
```

```
RESULT 11
LOCUS BC011062 2334 bp mRNA linear HTC 04-MAR-2003
DEFINITION Mus musculus, Similar to utrophin, clone IMAGE:3979320, mRNA.
ACCESSION BC011062
VERSION BC011062.1 GI:15029695
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (Bases 1 to 2334)
```

AUTHORS

Strausberg, R.
Direct Submission
Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: angbcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulesged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 23 Row: j Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein
this clone has the following problem: retained intron.

Location/Qualifiers

1. .2334

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3979320"

/issue_type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."

/clone_lib="NCI CGAP_Mam1"

/lab_host="DH10B"

/notes="Vector: pCMV-SPORT6"

BASE COUNT 698 a 512 c 638 g 486 t

ORIGIN

Query Match 16.3%; Score 343; DB 11; Length 2334;

Best Local Similarity 59.0%; Pred. No. 5e-58;

Matches 589; Conservative 0; Mismatches 410; Indels 0; Gaps 0;

Qy 104 AGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGTA 163

Db 1042 AGCACCGTCACTGAAGTGGACATGGATTGGACAGCTACAGATCCGCTAGAGGATG 1101

Qy 164 TTATCGTGGCTTCTTCTGCTGAGGACACATTCGCAAGCAAGGAGAGATTTCTAATGAT 223

Db 1102 CTGACGTGGCTGTGTCGCGGAGGACAGCTCCAGGAGCAAGATGACATTTCTGATGAT 1161

Qy 224 GTGGAAGTGTGAAAGACCACTTTCTATCTCATGAGGGGTACATGATGATTTGACAGCC 283

Db 1162 GTCGAAGAGTCAAGAGCAGGATTTGTCATCCCATGAAACTTTTATGATGAGCTGACGCA 1221

Qy 284 CATTGAGGCGCGGTGGTATATTTCTACATTTGGGAAGTAAGCTGATTTGGAAACAGGAAA 343

Db 1222 CACGAGAGCAGCTGGGGAGCGTCTTCAGGCTGGCAACCGCTGATGACACAAGGACT 1281

Qy 344 TTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGATCTCTTAATCAAGATGG 403

Db 1282 CTGTGAGAGGAGGAGGATTTGAGATCCAGGAACAGATGACCTTGTCTGAATGCAAGGTGG 1341

Qy 404 GAATGCTCTCAGGCTAGCTAGCATGGGAAAACAAAGCAATTTACATAGAGTTTAAATGGAT 463

Db 1342 GAGCGCTCCGGTGGAGAGCATGGAGGCGACTCCCGCTGCACGCTCTGTGTGGAG 1401

Qy 464 CTCAGAAATCAGAAATCGAAGAGTTGAATGACTGTGCTAACAAAAACAGAAAGAAC 523

1402 CTGCAGAGAAACAGCTGCAGAGCTCTCAAGCTGGCTGGCCCTCAGAGAGCGCATT 1461
QY 524 AGGAATATGGAGAGAGCTCTTGGACCTGATCTTGAAGACCTAAACCCCAAGTACAA 583
Db 1462 CAGAGATGGAGAGCTCCCGCTGGGTGATGACCTGCGCTCCCTCGAGAGCTGCTTCAA 1521
QY 584 CAACATAGGTGCTTCAAGAGATCTAGAACAAAGAACAGTCAAGTCAATCTCTCACT 643
Db 1522 GAACATAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGTAAATTCCTTAAT 1581
QY 644 CACATGGTGGTGTAGTATGATATCTAGTGGAGATCACCAACTGCTGCTTGGAGAA 703
Db 1582 CACATGGTGGTGTAGTATGATATCTAGTGGAGATCACCAACTGCTGCTTGGAGAT 1641
QY 704 CAACCTTAAGTATTTGGGAGATCGATGGCGAAACATCTGTAGATGGAAGAGACCGCTGG 763
Db 1642 CAGTTACAGAACTGGGTGAGCGCTGGACAGCTGATGCGCGTGGACTGAAGAACGTTGG 1701
QY 764 GTTCTTTTACAGACATCTCTCTCAATATGGCAAGCTCTTACTGAGAACAGTGCCTTTT 823
Db 1702 AACAGTTTGAAGAAATCAGTATTTCTGTGGCAGGAATTTATGGAAGAGCAGTGTCTGTTG 1761
QY 824 AGTGCATGCTTTTCAAGAAAGAGATGCTAGTGAACAAAGATTCACAACTGGCTTTAAA 883
Db 1762 GAGGCTTGGCTCACCGAAGAGAGAGCTTTGATTAAGTTCAACACGCAACTTTAAA 1821
QY 884 GATCAAAATGAATGTTATCAAGTCTTCAAAACCTGGCCGTTTAAAGCGGATCTAGAA 943
Db 1822 GACCAGAGAGAACTAAGTGTCAAGTGTCCGCGCTGCTGCTATATTAAGAGAGACATGAA 1881
QY 944 AAGAAAGAGATCCATGGCAACTGATTTCACTCAACAGAGATCTTCTTCAACACTG 1003
Db 1882 ATGAAGAGCGACCTCTGGATCACTGATGAGATGGCCAGGATGTGGCCCAATCTC 1941
QY 1004 AAGAATAAGTCAAGTCAAGCAGAGAGAGAGAGATGGCTGGATGATCTTGGCCGCTGTTG 1063
Db 1942 AGTAATCCCAAGGATCTAAGAGATGAACAGTCACTCTGAGGAGCTAACAACAGAGATG 2001
QY 1064 GATAATTTAGTCCAAACTTGAAGAGATGACAGACAG 1102
Db 2002 GATTCCTGGTTCAGAGACTCGAAGACTCTTCTTAACCAAG 2040

RESULT 12
LOCUS CB547284/c 402 bp mRNA linear EST 01-APR-2003
DEFINITION AMGNNUC:SRPB2-00242-E3-A srpb2 (10220) Rattus norvegicus cDNA clone
srpb2-00242-e3 5', mRNA sequence.

ACCESSION CB547284
VERSION CB547284.1 GI:29431225
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 402)
AUTHORS Angen EST Program.
TITLE Angen Rat EST Program
JOURNAL Unpublished
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00242 row: e column: 3.
High quality sequence stop: 3.
High quality sequence stop: 771.

FEATURES
source
1. .402
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="srpb2-00242-e3"
/tissue type="prostate tissue"
/clone_lib="srpb2 (10220)"

/note=Vector: pSPORT1; Site_1: SalI; Site_2: NotI; rat
prostate normalized double selected poly(A+) mRNA size
fraction > 1 kb"
BASE COUNT 83 a 111 c 88 g 119 t 1 others
ORIGIN

Query Match 15.4%; Score 324.6; DB 14; Length 402;
Best Local Similarity 90.8%; Pred. No. 1.9e-54;
Matches 345; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1519 TGATGAATTAAGCCCGCAGGACCTATTGGAGGGGAGCTTTCCAGCAGTTCAAGAGCAGAA 1578
Db 402 TGATCAGTTGAGCCGCGCAGGACCTATTGGTGGCGATTTTCCAGCAGTTCAAGAGCAGAA 343
QY 1579 COATGTAATAGGGGCTTCAAGAGGGAATTAAGAACTAAGAACTGTAATCATGAGTAC 1638
Db 342 TGATGTACACAGGGGCTTCAAGAGGGAATTAAGAACTAAGAACTGTAATCATGAGTAC 283
QY 1639 TCTTCAGACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTTAGAGAACTCTA 1698
Db 282 TCTGAGACTGTGAGAAATTTCTGACAGAGCAGCCTTTTGAAGGACTTAGAGAACTCTA 223
QY 1699 CCAGAGCCCGCAGAGAGCTGCTCTCTGAGAGAGAGCCCGAGAAATGTCACTCGGCTTCTACG 1758
Db 222 CCAGAGCCCGCAGAGAGCTGCTCTCTGAGAGAGAGCTCAGAAATGTCACTCGGCTTCTACG 163
QY 1759 AAGCAGGCTGAGAGAGTCACTGAGTGGGAAAATTTGAACCTGCACTCCGCTGACTG 1818
Db 162 AAGCAGGCTGAGAGAGTCACTGAGTGGGAAAATTTGAACCTGCACTCCGCTGACTG 103
QY 1819 CGAGAGAAAATAGATGAGAGCCCTTTGAAGAGCTCCAGAACTTCAAGAGGCCAGGATGA 1878
Db 102 CGAGCGAAAATAGATGAGAGCTTTGAAGAGCTCCAGAACTTCAAGAGGCCAGGATGA 43
QY 1879 GCTGGACCTCAAGCTCGGCC 1898
Db 42 ACAGCAGCTCATGTTGGCGC 23

RESULT 13
LOCUS BI250598 772 bp mRNA linear EST 17-JUL-2001
DEFINITION 602993659F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5149431 5',
mRNA sequence.

ACCESSION BI250598
VERSION BI250598.1 GI:14799101
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 772)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabp@remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLAM11369 row: e column: 16
High quality sequence start: 3
High quality sequence stop: 771.

FEATURES
source
1. .772
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

BU201022 851 bp mRNA linear EST 25-NOV-2002 241
LOCUS 603952191F1 CSEQHN03 Gallus gallus CDNA clone CHEST914e23 5', mRNA 646
DEFINITION sequence.
ACCESSION BU201022
VERSION BU201022.1 GI:25364517
KEYWORDS EST.
SOURCE Gallus gallus (chicken) 301
ORGANISM Gallus gallus 706
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; 766
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; 421
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 851)
AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1..851
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST914e23"
/tissue type="whole embryo"
/dev stage="20-21"
/lab_host="DH10B"
/clone_lib="CSEQHN03"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was bluntended, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91:9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
BASE COUNT 269 a 169 c 238 g 175 t
ORIGIN

Query Match 10.3%; Score 217.2; DB 13; Length 851;
Best Local Similarity 51.8%; Pred. No. 5.4e-33;
Matches 362; Conservative 0; Mismatches 223; Indels 1; Gaps 1;
347 TCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGGAA 406
2 TCATCTCAGAGGAGTTGAGATTAGAAACAAATGCTGCTGAACCTCCGCTGGGAG 61
407 TGCCTCAGGCTAGTACATGGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTC 466
62 GACCTCAGGCTGAGAGCATGGAGCAGCTCCGCTGCTGACATGCTGATGGAGCTA 121
467 CAGAAATCAGAAATCGAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAACAAAG 526
122 CAAAGAGCAGTTGACAGAGCTGCTGCTGAGCTGAGCTACAGAGAACGCAATTCAG 181
527 AAAATGGAGGAGAGGCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTACACAA 586

Db 182 AAGATGGAATCTCAACTCTTAGCTGAAGATTTAGAAGCCCTTCAAAAACAGCTAGAGAA 241
QY 587 CATAGGTGCTTCAAGAGAGATCTAGAACAGAACAGAGTCAGGGTCAATTTCTCTCACTCAC 646
Db 242 CACAAAAGCTTGCAAAGTGACCTGGAGCGCAACAGTGAAGGTAACTCTCTGCGCAC 301
QY 647 ATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAACAA 706
Db 302 ATGGTTGTCATCGTTGATGAAAGCAGTGGGGAGAGTGCACACAGCTGTTTCTGGAGAACAA 361
QY 707 CTTAAGGTATTGGGAGATCGATGGGCAACACATCTGTAGATGAGACAGAGACCGCTGGGTT 766
Db 362 TTGCGAAGCTGGGTGGAGCGGTGGACAGCAGTTTGTCTGACATGAGGAAGACGATC 421
QY 767 CTTTACAGACATCTTCTCAATGGCAAGCTCTTACTGAAGAACAGCTGCTTTTAACT 826
Db 422 AAATTGCAAGAAATTCAGCTTCTATGGCAGGAGCTGTAGAAGACGACGTGCTTATTGAAG 481
QY 827 GCATGGCTTTTCAAAAAGAGATGACAGTGAACAGATTTCACACAACTGGCTTTAAAGAT 886
Db 482 GCGTGGTTAACTGAGAAGGAAGATGCTTAAGTAAGTCCAGACCAAGCACTTCAGAGAC 541
QY 887 CAAATGAAATG-TTATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAA 931
Db 542 CAAATGAGATCTGAGTGTGAACCTGTTGGAACCTAGCAATTTTGA 587

Search completed: February 2, 2004, 06:03:04
Job time : 3948.93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
M nucleic - nucleic search, using sw model
run on: February 1, 2004, 23:51:42 ; Search time 697.863 Seconds
(without alignments)
1131.886 Million cell updates/sec
File: US-09-845-416-2_COPY_900_3000
Perfect score: 2101
Sequence: 1 gacgtatgctacacagg.....acctcagcactctggaagac 2101
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2434939 seqs, 1822278265 residues 4869878
Total number of hits satisfying chosen parameters:
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2101	100.0	4182	13	US-09-845-416-2 Sequence 2, Appl
2	2101	100.0	5149	13	US-09-845-416-27 Sequence 27, Appl
3	1579.8	75.2	3999	13	US-09-845-416-6 Sequence 6, Appl
4	1579.8	75.2	4966	13	US-09-845-416-28 Sequence 28, Appl
5	1579.8	75.2	4990	13	US-09-845-416-34 Sequence 34, Appl
6	1443	68.7	3858	13	US-09-845-416-9 Sequence 9, Appl
7	1443	68.7	4825	13	US-09-845-416-29 Sequence 29, Appl
8	1443	68.7	4848	13	US-09-845-416-35 Sequence 35, Appl
9	1443	68.7	5060	13	US-09-845-416-36 Sequence 36, Appl
10	1293	61.5	8689	13	US-10-149-736-42 Sequence 42, Appl
11	1114	53.0	4414	13	US-09-845-416-32 Sequence 32, Appl
12	1103.4	52.5	5417	13	US-10-149-736-39 Sequence 39, Appl
13	1103.4	52.5	11443	13	US-10-149-736-44 Sequence 44, Appl
14	1103.4	52.5	12057	13	US-10-149-736-47 Sequence 47, Appl
15	1103.4	52.5	13957	10	US-09-782-378A-22 Sequence 22, Appl

16	1103.4	52.5	13957	10	US-09-880-107-2284 Sequence 2284, Ap
17	1103.4	52.5	13957	13	US-10-149-736-1 Sequence 1, Appl
18	1103.4	52.5	14082	13	US-10-341-434-108 Sequence 108, App
19	1102	52.5	3446	13	US-09-845-416-14 Sequence 14, Appl
20	1100.2	52.4	11058	13	US-09-845-416-1 Sequence 1, Appl
21	1097.8	52.3	1951	13	US-09-845-416-3 Sequence 3, Appl
22	999	47.5	2169	13	US-09-845-416-4 Sequence 4, Appl
23	999	47.5	3531	13	US-09-845-416-10 Sequence 10, Appl
24	999	47.5	4498	13	US-09-845-416-30 Sequence 30, Appl
25	997	47.5	5339	13	US-10-149-736-40 Sequence 40, Appl
26	996	47.4	5462	13	US-10-149-736-41 Sequence 41, Appl
27	866.6	41.2	13815	13	US-10-149-736-2 Sequence 2, Appl
28	787	37.5	3510	13	US-09-845-416-12 Sequence 12, Appl
29	777	37.0	4476	13	US-09-845-416-31 Sequence 31, Appl
30	777	37.0	1667	13	US-09-845-416-7 Sequence 7, Appl
31	652	31.0	1821	13	US-09-845-416-13 Sequence 13, Appl
32	450	21.4	1340	13	US-09-845-416-11 Sequence 11, Appl
33	387	18.4	387	13	US-10-149-736-32 Sequence 32, Appl
34	361.4	17.2	10302	10	US-09-782-378A-23 Sequence 23, Appl
35	361.4	17.2	10302	13	US-10-149-736-3 Sequence 3, Appl
36	361.4	17.2	16531	13	US-10-101-510-667 Sequence 667, App
37	348	16.6	348	13	US-10-149-736-31 Sequence 31, Appl
38	336.6	16.0	11096	13	US-10-149-736-4 Sequence 4, Appl
39	333	15.8	333	13	US-10-149-736-9 Sequence 9, Appl
40	327	15.5	327	13	US-10-149-736-8 Sequence 8, Appl
41	322.4	15.3	333	13	US-10-149-736-10 Sequence 10, Appl
42	265	12.6	1434	13	US-09-845-416-15 Sequence 15, Appl
43	261	12.4	324	13	US-10-149-736-33 Sequence 33, Appl
44	190.2	9.1	256	9	US-09-864-761-21956 Sequence 21956, A
45	180	8.6	476	9	US-09-864-761-15766 Sequence 15766, A

ALIGNMENTS

RESULT 1
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match	100.0%	Score	2101;	DB	13;	Length	4182;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	2101;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	GAGCTATGCTTACACACAGGCTGCTTATGTCCACACCTCTGACCCCTACACGAGGCCATT	60				
DB	900	GAGCTATGCTTACACACAGGCTGCTTATGTCCACACCTCTGACCCCTACACGAGGCCATT	959				
QY	61	TCCTTTCACAGCATTGGAAAGCTCTCTGAAGACAAAGTCATTGGCAGTTCATTGATGGAGAG	120				
DB	960	TCCTTTCACAGCATTGGAAAGCTCTCTGAAGACAAAGTCATTGGCAGTTCATTGATGGAGAG	1019				
QY	121	TGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAGAGTATATCGTGGCTTTCTTTC	180				
DB	1020	TGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAGAGTATATCGTGGCTTTCTTTC	1079				
QY	181	TGCTGAGGACACATTGCAAGCAGCAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA	240				

SOFTWARE: PatentIn Ver. 2.1.1

SEQ ID NO 27

LENGTH: 5149

TYPE: DNA

ORGANISM: Homo sapiens

IS-09-845-416-27

Query Match 100.0%; Score 2101; DB 13; Length 5149;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2557	ATCAAGTCTTCAAAACTGGCCGTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	2616
961	GGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACAACCTGAAGAATAAGTCAGTGAC	1020
2617	GGCAAACTGTATTCACTCAAAACAAGATCTTCTTTCAACAACCTGAAGAATAAGTCAGTGAC	2676
1021	CCAGAAAGCGAAGCATGGCTGGATTAACCTTTGCCCGGTCTTGGGATAATTTAGTCCAAA	1080
2677	CCAGAAAGCGAAGCATGGCTGGATTAACCTTTGCCCGGTCTTGGGATAATTTAGTCCAAA	2736
1081	ACTTTGAAAAGAGTACAGCAGACTCATAGATTACTGCAACAGATTTCCCTCGACCTGGGA	1140
2737	ACTTTGAAAAGAGTACAGCAGACTCATAGATTACTGCAACAGATTTCCCTCGACCTGGGA	2796
1141	AAAGTTTCTTGGCTGTACAGAACTGAAACAACTGCGAATGTCCTACAGAGATGCTAC	1200
2797	AAAGTTTCTTGGCTGTACAGAACTGAAACAACTGCGAATGTCCTACAGAGATGCTAC	2856
1201	CCGTAAAGAAAGCTCCTAGAAAGCTCCAAAGGAGTAAAAGAGCTGATGAAAACAAATGGCA	1260
2857	CCGTAAAGAAAGCTCCTAGAAAGCTCCAAAGGAGTAAAAGAGCTGATGAAAACAAATGGCA	2916
1261	AGACCTCCAAAGGTGAAATTTGAAGCTCACACAGATGTTATCAACCTGGATGAAAACAG	1320
2917	AGACCTCCAAAGGTGAAATTTGAAGCTCACACAGATGTTATCAACCTGGATGAAAACAG	2976
1321	CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCCTGTACAAAAGAGCTTT	1380
2977	CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCCTGTGTCAAAAGAGCTTT	3036
1381	GGATAACATCAACTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAAGTCCCA	1440
3037	GGATAACATCAACTTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAAGTCCCA	3096
1441	TTTGGAAAGCCAGTCTGACAGTGGAGGCTGTGACCTTTCTCTGAGGAACTTCTGGT	1500
3097	TTTGGAAAGCCAGTCTGACAGTGGAGGCTGTGACCTTTCTCTGAGGAACTTCTGGT	3156
1501	GTGGCTACAGCTCAAGAGATGATCAATTTAAGCCGGCAGGACCTATTGGAGGCGACTTTCC	1560
3157	GTGGCTACAGCTCAAGAGATGATCAATTTAAGCCGGCAGGACCTATTGGAGGCGACTTTCC	3216
1561	AGCAGTTTCAAGACAGAAACGATGTATAGGGCTTCAAGAGGGAATTTGAAAATAAGA	1620
3217	AGCAGTTTCAAGACAGAAACGATGTATAGGGCTTCAAGAGGGAATTTGAAAATAAGA	3276
1621	ACCTGTAATCATGAGTACTCTTCAGACTGTACGAATATTTCTCAGACAGAGCCTTTGGA	1680
3277	ACCTGTAATCATGAGTACTCTTCAGACTGTACGAATATTTCTCAGACAGAGCCTTTGGA	3336
1681	AGGACTAGAGAAAACCTTACCCAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGCCCGAGAA	1740
3337	AGGACTAGAGAAAACCTTACCCAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGCCCGAGAA	3396
1741	TGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAAATTTGAA	1800
3397	TGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAAATTTGAA	3456
1801	CCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAAGACTCCAGGAAT	1860
3457	CCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAAGACTCCAGGAAT	3516
1861	TCAAGAGGCCACCGATGAGCTGGAACCTTCAAGCTGCGCAAGCTGAGTGATCAAGGATC	1920
3517	TCAAGAGGCCACCGATGAGCTGGAACCTTCAAGCTGCGCAAGCTGAGTGATCAAGGATC	3576
1921	CTGGCAGCCGCTGGGCGATCTCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAA	1980
3577	CTGGCAGCCGCTGGGCGATCTCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAA	3636
1981	GGCACTTCGAGGAGAAAATTGGCTCTTGAAGAGAAAGCTGAGCCACGTCATGACCTTGC	2040
3637	GGCACTTCGAGGAGAAAATTGGCTCTTGAAGAGAAAGCTGAGCCACGTCATGACCTTGC	3696

QY 2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAGCTCTGGAAGA 2100
Db 3697 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAGCTCTGGAAGA 3756
QY 2101 C 2101
Db 3757 C 3757

RESULT 3
US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-6

Query Match 75.2%; Score 1579.8; DB 13; Length 3999;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

QY 1 GAGCTATGCTTACACAGGCTGTATGTACACCACTCTGACCCCTACAGGAGCCCAATT 60
Db 900 GAGCTATGCTTACACAGGCTGTATGTACACCACTCTGACCCCTACAGGAGCCCAATT 999
QY 61 TCGTTTACACAGCATTTGGAGGCTCTTGAAGAGCAAGTCAATTTGGCAGTTCAATGTAGAGAG 120
Db 960 TCGTTTACACAGCATTTGGAGGCTCTTGAAGAGCAAGTCAATTTGGCAGTTCAATGTAGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAGAGTATATCGTGGCTTCTTTC 180
Db 1020 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAGAGTATATCGTGGCTTCTTTC 1079
QY 181 TCGTAGGACACATTCGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTCAAGA 240
Db 1080 TCGTAGGACACATTCGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTCAAGA 1139
QY 241 CCAGTTTCACTCATAGGGGTACATGATGATTTGACGCCCATCAGGCCCATCAGGCCCGGTGG 300
Db 1140 CCAGTTTCACTCATAGGGGTACATGATGATTTGACGCCCATCAGGCCCATCAGGCCCGGTGG 1199
QY 301 TAATATTTACAAATTTGGAGTAAAGTGAAGTGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1200 TAATATTTACAAATTTGGAGTAAAGTGAAGTGAACAGGAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAAGTACAGAGCAGATGAATCTCCTAATTCAGATGGGAATGCTCAGGGTAGC 420
Db 1260 AACTGAAGTACAGAGCAGATGAATCTCCTAATTCAGATGGGAATGCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAACCAAGCAATTTACATAGATTTTATGGAATCTCCAGAATCAGAACT 480
Db 1320 TAGCATGGAAAACCAAGCAATTTACATAGATTTTATGGAATCTCCAGAATCAGAACT 1379
QY 481 GAAAGAGTTGAATCACTGGCTAAACAAAACAGAGAAAGAACAGGAAATTTGAGGAGAGA 540
Db 1380 GAAAGAGTTGAATCACTGGCTAAACAAAACAGAGAAAGAACAGGAAATTTGAGGAGAGA 1439
QY 541 GCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACACACATTAAGTGCCTTCA 600

Db 1440 GCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAGGTGCTTCA 1499
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACAATGCTGTGTAGT 660
Db 1500 AGAAGATCTAGAACAAAGAACAAAGTCAGGTCATTTCTCTCACTCACAATGCTGTGTAGT 1559
QY 661 TGATGAATCTAGTGGAGATCACGCAATCTGCTGTGTTGGAGAGAACAACTTAAGGTATTGGG 720
Db 1560 TGATGAATCTAGTGGAGATCACGCAATCTGCTGTGTTGGAGAGAACAACTTAAGGTATTGGG 1619
QY 721 AGATCGATGGGCAACATCTCTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 1620 AGATCGATGGGCAACATCTCTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACCA 1679
QY 781 CTTCTCTCAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTTACA 840
Db 1680 GCCTGAC-----CTAGCTCTCTGGACTCACCACATATTGGAGCCTCTCTACTCA 1727
QY 841 AAAGAAGATGCGAGTGAACAAGATTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1728 GACTGTTACTCTGGTGACACA-----ACCTGTGGTTACTTAAGGAACTGCCAATCT- 1777
QY 901 ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1778 ----- 1777
QY 961 GGGCAAACTGATTCATCTCAACAAGATCTTTTCAACACTGAGATTAAGTCAGTGAC 1020
Db 1778 -----CCAAACTAGAAAATGCCATCTTCTCTGATTTGGAG----- 1812
QY 1021 CCAGAAGACGGAAGCATGGCTGGATACTTTGCCCGGTGTTGGGATATTTAGTCCAAAA 1080
Db 1813 ----- 1812
QY 1081 ACTTGAAAAGAGTACAGCAGACAGCTCATAGATTACTGCAACAGTTCCCTCCCTGGACTGA 1140
Db 1813 -----GTACCTACTCATAGATTACTGCAACAGTTCCCTCCCTGGACTGA 1856
QY 1141 AAAGTTTCTTGCTGGCTTACAGAACTGAAACAACCTGCCAATGCTCTACAGAGTGTCTAC 1200
Db 1857 AAAGTTTCTTGCTGGCTTACAGAACTGAAACAACCTGCCAATGCTCTACAGAGTGTCTAC 1916
QY 1201 CCGTAGGAAAAGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGTAGAAACAATGGA 1260
Db 1917 CCGTAGGAAAAGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGTAGAAACAATGGA 1976
QY 1261 AGACTCCAGAGTGAATTTGAAGCTTCAACAGATGTTTATCAACCTGGATGAAAACAG 1320
Db 1977 AGACTCCAGAGTGAATTTGAAGCTTCAACAGATGTTTATCAACCTGGATGAAAACAG 2036
QY 1321 CCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGATGATGATGATGATGATGATGAT 1380
Db 2037 CCAAAAAATCCTGAGATCCCTGGAAAGTTCCGATGATGATGATGATGATGATGATGATGAT 2096
QY 1381 GATTAACATGAATTCAGTGGAGTGAATCTCGGAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 2097 GATTAACATGAATTCAGTGGAGTGAATCTCGGAAAAGTCTCTCAACATTAGGTCCCA 2156
QY 1441 TTTGGAAGCCAGTTCTGACCACTGGAAGGCTGTGACCTTTCTCTGACGAACTTCTGT 1500
Db 2157 TTTGGAAGCCAGTTCTGACCACTGGAAGGCTGTGACCTTTCTCTGACGAACTTCTGT 2216
QY 1501 GTGGCTACAGCTGAAGATGAATTAAGCCGGCAGGACCTATTTGGAGGCGACTTCC 1560
Db 2217 GTGGCTACAGCTGAAGATGAATTAAGCCGGCAGGACCTATTTGGAGGCGACTTCC 2276
QY 1561 AGCAGTTTCAAGACAGAACGATGTACATAGGGCTTTCAAGAGGGAATTTGAAACTAAAGA 1620
Db 2277 AGCAGTTTCAAGACAGAACGATGTACATAGGGCTTTCAAGAGGGAATTTGAAACTAAAGA 2336
QY 1621 ACCTGTAATCATGAGTACTCTTGAAGCTGTACGAATATTTCTGACAGACGCGCTTGA 1680
Db 2337 ACCTGTAATCATGAGTACTCTTGAAGCTGTACGAATATTTCTGACAGACGCGCTTGA 2396

```
1681 AGGACTAGAGAACTCTACCGAGGCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAA 1740
1682 AGGACTAGAGAACTCTACCGAGGCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAA 2456
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATCTAGTGGGAAAAATTGAA 1800
2457 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATCTAGTGGGAAAAATTGAA 2516
1801 CTTGCACTCGGCTGAGTGCAGAGAAATATAGATGAGACCTTTGAAAGACTCCAGGAACT 1860
2517 CTTGCACTCGGCTGAGTGCAGAGAAATATAGATGAGACCTTTGAAAGACTCCAGGAACT 2576
1861 TCAAGAGGCCAGGATGAGTGCACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 1920
2577 TCAAGAGGCCAGGATGAGTGCACCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGGATC 2636
1921 CTGGAGCCCGTGGGGATCTCTCTATGACTCTCTCCNAGTCACTTCGAGAAAGTCAA 1980
2637 CTGGAGCCCGTGGGGATCTCTCTATGACTCTCTCCNAGTCACTTCGAGAAAGTCAA 2696
1981 GGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAACTGAGCCACGTCAATGACCTTGC 2040
2697 GGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAACTGAGCCACGTCAATGACCTTGC 2756
2041 TGCCAGCTTACACCTTTGGGCAATTCAGTCTCTCAACGATTAACCTCAGCACTCTGGAAGA 2100
2757 TGCCAGCTTACACCTTTGGGCAATTCAGTCTCTCAACGATTAACCTCAGCACTCTGGAAGA 2816
2101 C 2101
2817 C 2817
```

RESULT 4
JS-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 75.2%; Score 1579.8; DB 13; Length 4966;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

```
2Y 1 GAGCTATGCTACACACAGGCTGTTATGTCACACCTCTGACCCCTACACGGAGCCCAATT 60
DB 1657 GAGCTATGCTACACACAGGCTGTTATGTCACACCTCTGACCCCTACACGGAGCCCAATT 1716
QY 61 TCCTTCACAGCAATTTGAGCTCTGAGACAACTCATTTGGCAGTTCATTTGATGAGAG 120
DB 1717 TCCTTCACAGCAATTTGAGCTCTGAGACAACTCATTTGGCAGTTCATTTGATGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGTTCATCAACAGCTTTAGAGAGATTAATCGTGGCTTTCTTC 180
DB 1777 TGAAGTAAACCTGGACCGTTCATCAACAGCTTTAGAGAGATTAATCGTGGCTTTCTTC 1836
QY 181 TGCTGAGACACATTTGACGACACAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
```

```
DB 1837 TGCTGAGGACACATTTGCAAGCAACAGGAGAGANTTTCTAATGATGTGGAAGTGGTGAAAGA 1896
QY 241 CCAGTTTCTACTCATGAGGGGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 300
DB 1897 CCAGTTTCTACTCATGAGGGGTACATGATGATGATGATGATGATGATGATGATGATGATG 1956
QY 301 TAAATATTTACAAATTTGGGAAGTAAAGCTGATTTGGAACAGGAAAAATTTATCAGAGATGAAGA 360
DB 1957 TAAATATTTACAAATTTGGGAAGTAAAGCTGATTTGGAACAGGAAAAATTTATCAGAGATGAAGA 2016
QY 361 AACTGAAGTACAAAGACAGATGAATCTCTAAATTCAGATGGAATGCTTCAGGGTAGC 420
DB 2017 AACTGAAGTACAAAGACAGATGAATCTCTAAATTCAGATGGAATGCTTCAGGGTAGC 2076
QY 421 TAGCATGGAAAAACAAAGCAATTTATATAGAGTTTAAATGATGATGATGATGATGATGATG 480
DB 2077 TAGCATGGAAAAACAAAGCAATTTATATAGAGTTTAAATGATGATGATGATGATGATGATG 2136
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGAAAAATGGAGGAAGA 540
DB 2137 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAAGAAACAAAGAAAAATGGAGGAAGA 2196
QY 541 GCCTCTTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTACAAACAAATAGGTGCTTCA 600
DB 2197 GCCTCTTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTACAAACAAATAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAACAGAAACAGTTCAGGGTCAATTTCTCTCCTCAGTCAATGATGATGATG 660
DB 2257 AGAAGATCTAGAACAGAAACAGTTCAGGGTCAATTTCTCTCCTCAGTCAATGATGATGATG 2316
QY 720 TGATGAATCTTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACTTAAGTATTTGGG 720
DB 2317 TGATGAATCTTAGTGGAGATCACGCAACTGCTGCTTTTGGAGAAACAACTTAAGTATTTGGG 2376
QY 721 AGATCGATGGGCAAAACATCTGTAGATGAGACAGAAAGCCGCTGGGTTCTTTTACAAGACAT 780
DB 2377 AGATCGATGGGCAAAACATCTGTAGATGAGACAGAAAGCCGCTGGGTTCTTTTACAAGACCA 2436
QY 781 CCTTCTCAATGGCAACGCTCTTACTGAAGAAACAGTGCCTTTTGTAGTGCATGGGTTTCAGA 840
DB 2437 GCCTGAC-----CTAGCTCTGAGCTGACCACTATTGGAGCCCTCTCCTACTCA 2484
QY 841 AAAAGAGATGCAAGTAAACAGATTCACAACTGGCTTTAAAGATCAAAATGAATGTT 900
DB 2485 GACTGTTACTCTGCTGACACA-----ACCTGTGGTTACTTAAGGAAACTGCCATCT- 2534
QY 901 ATCAAGTCTTCAAAAACCTGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
DB 2535 ----- 2534
QY 961 GGGCAAACTGTATTCACTCAAAACAGATCTTTCTTCAACACTGAAGAAATAGTCAAGTAC 1020
DB 2535 -----CCAAACTAGAAATGCCATCTTCTCTGATGTTGGAG----- 2569
QY 1021 CCAGAGACGGAAGCATGGCTGGATTAACCTTTGCCGCTGTGGGATAATTTAGTCCAAA 1080
DB 2570 ----- 2569
QY 1081 ACTTGAAGAGTACAGACAGACTCATAGATTACTGCAACAGTTCCCTCGGACCTTGA 1140
DB 2570 -----GTAACCTACTCATAGATTACTGCAACAGTTCCCTCGGACCTTGA 2613
QY 1141 AAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCGCAATGTCTCTACAGATGCTAC 1200
DB 2614 AAAGTTTCTTGCCTGGCTTACAGAGCTGAAACAACTGCGCAATGTCTCTACAGATGCTAC 2673
QY 1201 CCGTAAAGGAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAAATGGCA 1260
DB 2674 CCGTAAAGGAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAAATGGCA 2733
QY 1261 AGACCTCCAAAGGTGAAATTTGAAGCTCAGAGATGTTTATCAACAACTGGATGAAACAG 1320
DB 2734 AGACCTCCAAAGGTGAAATTTGAAGCTCAGAGATGTTTATCAACAACTGGATGAAACAG 2793
```


1321 CCAAAAATCCTGAGATCCCTGGAGGTTCCGATGATCGAGTCTGTTTACAAAGACGTTT 1380
Db CCAAAAATCCTGAGATCCCTGGAGGTTCCGATGATCGAGTCTGTTTACAAAGACGTTT 2853
1381 GGATAACATGAACCTCAAGTCGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCA 1440
Db GGATAACATGAACCTCAAGTCGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGGTCCCA 2913
1441 TTGGAAGCCAGTTCTGACCAATGGAAGCGTCTGACCTTCTCTGAGGAACTCTGGT 1500
Db TTGGAAGCCAGTTCTGACCAATGGAAGCGTCTGACCTTCTCTGAGGAACTCTGGT 2973
1501 GTGGCTACAGCTGGAAGATGATGAATTAAGCCGCGAGGCACTTATGAGGCGCACTTTC 1560
Db GTGGCTACAGCTGGAAGATGATGAATTAAGCCGCGAGGCACTTATGAGGCGCACTTTC 3033
1561 AGCAGTTTCAAGAGCAAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db AGCAGTTTCAAGAGCAAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 3093
1621 ACCTGTAATCATGACTACTCTTTGAGACTGTACGATATTTCTGACAGAGAGCCCTTTTGA 1680
Db ACCTGTAATCATGACTACTCTTTGAGACTGTACGATATTTCTGACAGAGAGCCCTTTTGA 3153
1681 AGGACTAGAGAACTCTACAGGAGCCGAGAGAGTGTGCTCTCTGAGGAGAGAGCCCGAGAA 1740
Db AGGACTAGAGAACTCTACAGGAGCCGAGAGAGTGTGCTCTCTGAGGAGAGAGCCCGAGAA 3213
1741 TGTCACTCGGCTTACGAAAGCAGGCTGAGGAGTCAATGATGAGTGGGAAAAATTGAA 1800
Db TGTCACTCGGCTTACGAAAGCAGGCTGAGGAGTCAATGATGAGTGGGAAAAATTGAA 3273
1801 CTTGCACTCGGCTGAGTGGGAGTCAATGATGAGTGGGAAAAATTGAA 1860
Db CTTGCACTCGGCTGAGTGGGAGTCAATGATGAGTGGGAAAAATTGAA 3333
1861 TCAAGAGGCGCGAGTGGGAGTCAATGATGAGTGGGAAAAATTGAA 1920
Db TCAAGAGGCGCGAGTGGGAGTCAATGATGAGTGGGAAAAATTGAA 3393
1921 CTGGCAGCGCGTGGGAGTCAATGATGAGTGGGAAAAATTGAA 1980
Db CTGGCAGCGCGTGGGAGTCAATGATGAGTGGGAAAAATTGAA 3453
1981 GGCACCTTCGAGGAGAAATTGGGAGTCAATGATGAGTGGGAAAAATTGAA 2040
Db GGCACCTTCGAGGAGAAATTGGGAGTCAATGATGAGTGGGAAAAATTGAA 3513
2041 TCGCAGCTTACCACTTGGGAGTCAATGATGAGTGGGAAAAATTGAA 2100
Db TCGCAGCTTACCACTTGGGAGTCAATGATGAGTGGGAAAAATTGAA 3573
2101 C 2101
3574 C 3574

RESULT 5

US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-34
Query Match 75.2%; Score 1579.8; DB 13; Length 4990;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
1 GAGCTATGCTACACACAGAGGCTGCTTATGTCACACCTCTCTGACCTCTCACCGAGCCCAATT 60
Db GAGCTATGCTTACACACAGAGGCTGCTTATGTCACCACTCTCTGACCTCTCACCGAGCCCAATT 1740
61 TCCTTCACAGCAATTTGGAAGCTCTCTGGAAGCAAGTCAATTTGCGAGTTCATTTGATGGAGAG 120
Db TCCTTCACAGCAATTTGGAAGCTCTCTGGAAGCAAGTCAATTTGCGAGTTCATTTGATGGAGAG 1800
121 TGAAGTAAACCTTGGACCGTTATCAAAAGCTTTAGAGAGTATTTATCGTGGCTCTCTTTC 180
Db TGAAGTAAACCTTGGACCGTTATCAAAAGCTTTAGAGAGTATTTATCGTGGCTCTCTTTC 1860
1801 TGCCTGAGGACACATTCGAAAGCAAGGAGAGATTTCTAATGATGGAAGTGGTGAAGA 240
Db TGCCTGAGGACACATTCGAAAGCAAGGAGAGATTTCTAATGATGGAAGTGGTGAAGA 1920
241 CCAAGTTTCATCTCATGAGGGGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db CCAAGTTTCATCTCATGAGGGGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
1921 TAATATTTACAAATTTGGGAACTTAAGCTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 360
Db TAATATTTACAAATTTGGGAACTTAAGCTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 2040
361 AACTGAGTACAGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db AACTGAGTACAGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
421 TAGCATGGAACCAAGCAATTTACATAGAGTCTTAAATGATCTCCAGAAATCAGAAATCT 480
Db TAGCATGGAACCAAGCAATTTACATAGAGTCTTAAATGATCTCCAGAAATCAGAAATCT 2160
481 GAAAGATTTGAATGATCTGGTCAACAAAACAGAGAAAGCAAGGAAATCGAGGAAAGA 540
Db GAAAGATTTGAATGATCTGGTCAACAAAACAGAGAAAGCAAGGAAATCGAGGAAAGA 2220
541 GCCTCTTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTACCAACCAATAAGGTGCTTCA 600
Db GCCTCTTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTACCAACCAATAAGGTGCTTCA 2280
601 AGAAGATCTTAGAAACAAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTGTAGT 660
Db AGAAGATCTTAGAAACAAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTGTAGT 2340
661 TGATGATCTTAGTGGAGATCAAGCACTGCTGTTTGGAGAAACCACTTAAGGTATTGGG 720
Db TGATGATCTTAGTGGAGATCAAGCACTGCTGTTTGGAGAAACCACTTAAGGTATTGGG 2400
721 AGATCGATGGGCAACATCTGTAGATGGAAGACCGTGGTCTTTTACAGACAT 780
Db AGATCGATGGGCAACATCTGTAGATGGAAGACCGTGGTCTTTTACAGACAT 2460
781 CTTCTCAATGGCAACGCTCTTACTGAGAAACAGTGTCTTTTGTAGTGGTCTTTTGTAGTGGT 840
Db CTTCTCAATGGCAACGCTCTTACTGAGAAACAGTGTCTTTTGTAGTGGTCTTTTGTAGTGGT 2508
2461 GCCTGAC-----CTAGCTCTGAGCTGACCACTATTGGAGCCCTCTCTACTCA 2508
841 AAAAGAGATGCACTGAAACAGATTTCAACAAGTGGCTTTTAAAGATCAAAATGAATGTT 900
Db GACTGTTACTCTGTGTGACACA-----ACCTGTGGTTTACTAAGGAAACTGCCATCT- 2558
901 ATCAAGTCTTCAAAAACATGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db ----- 2558

QY	2041	TCGCCAGCTTTACCACTTTGGGCATTTGAGTCTTCAACGTTATTAACCTCAGCACTCTGGAAGA	2100
Db	3538	TCGCCAGCTTTACCACTTTGGGCATTTGAGTCTTCAACGTTATTAACCTCAGCACTCTGGAAGA	3597
QY	2101	C 2101	
Db	3598	C 3598	
RESULT 6			
US-09-845-416-9			
; Sequence 9, Application US/09845416			
; Publication No. US2003017132A1			
; GENERAL INFORMATION:			
; APPLICANT: XIAO, XIAO			
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE			
; TITLE OF INVENTION: THEREOF			
; FILE REFERENCE: DE1142			
; CURRENT APPLICATION NUMBER: US/09/845,416			
; CURRENT FILING DATE: 2001-04-30			
; PRIOR APPLICATION NUMBER: 60/200,777			
; PRIOR FILING DATE: 2000-04-28			
; NUMBER OF SEQ ID NOS: 36			
; SOFTWARE: Patent In Ver. 2.1			
; SEQ ID NO 9			
; LENGTH: 3858			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-845-416-9			
Query Match 69.7%; Score 1443; DB 13; Length 3858;			
Best Local Similarity 84.6%; Pred. No. 0;			
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;			
QY	1	GAGCTATGCCCTACACACAGGCTGCTTATGTCCACACCTCTGACCCCTACACGGAGCCCACTT	60
Db	900	GAGCTATGCCCTACACACAGGCTGCTTATGTCCACACCTCTGACCCCTACACGGAGCCCACTT	959
QY	61	TCCTTTCACAGCATTTGGAAGCTCCTCGAAGACAAGTCAATTTGGCAGTTTCATTTGATGGAGAG	120
Db	960	TCCTTTCACAGCATTTGGAAGCTCCTCGAAGACAAGTCAATTTGGCAGTTTCATTTGATGGAGAG	1019
QY	121	TGAAGTAACTCGACCGTTATCAACAGCTTTAGAGAGAGTATTCGTTGGCTTCTTTC	180
Db	1020	TGAAGTAACTCGACCGTTATCAACAGCTTTAGAGAGAGTATTCGTTGGCTTCTTTC	1079
QY	181	TGCTGAGACACATTTGCAAGCACAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	240
Db	1080	TGCTGAGACACATTTGCAAGCACAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1139
QY	241	CCAGTTTCATCTCATGAGGGGTCATGATGGATTTGACAGCCCATCAGGGCCGGGTGG	300
Db	1140	CCAGTTTCATCTCATGAGGGGTCATGATGGATTTGACAGCCCATCAGGGCCGGGTGG	1199
QY	301	TAATATTTCTCAATTTGGGAAGTAAAGTTCATTGGAACAGGAAAAATTTATCAGAAGATGAAGA	360
Db	1200	TAATATTTCTCAATTTGGGAAGTAAAGTTCATTGGAACAGGAAAAATTTATCAGAAGATGAAGA	1259
QY	361	AACTGAAGTACAAGACAGAGATGAATCTCTTAATTAAGATGGGAATGCTTCAGGGTAGC	420
Db	1260	AACTGAAGTACAAGACAGAGATGAATCTCTTAATTAAGATGGGAATGCTTCAGGGTAGC	1319
QY	421	TAGCATGAAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	480
Db	1320	TAGCATGAAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT	1379
QY	481	GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAACAGAGAAATGGAGGAAGA	540
Db	1380	GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAACAGAGAAATGGAGGAAGA	1439
QY	541	GCCTCTTGACCTGATCTTTGAAGACCTTAAACGCCAAGTCAACAACAATAAGGTCCTTCA	600
Db	1440	GCCTCTTGACCTGATCTTTGAAGACCTTAAACGCCAAGTCAACAACAATAAGGTCCTTCA	1499

961	GGGCAAACTGTATTCTCACTCAAAACAGATCTTTCTTTCAACACATGAAAGAAATGATCAGTGAC	1020
2559	-----CCAAACTAGAAATGCCATCTTCTCTGATGTTGGAG-----	2593
1021	CCAGNAGCGAAGCATGGCTGGATAACTTTTGCCCGTGTGGGATAAATTTAGTCCAAAA	1080
2594	-----	2593
1081	ACTTGAAAAGAGTACAGCAGACCTCATGATTACTGCAACAGTTTCCCTCTGGAAGCTGGA	1140
2594	-----GTACTCTCATAGATTACTGCAACAGTTTCCCTCTGGAAGCTGGA	2637
1141	AAAGTTTCTTCCCTGGCTTACAGAAAGCTGAAACAACTGCCAAATGTCCTACAGAGTCTAC	1200
2638	AAAGTTTCTTCCCTGGCTTACAGAAAGCTGAAACAACTGCCAAATGTCCTACAGAGTCTAC	2697
1201	CCGTAAGAAAAGGCTCTTAGAAGACTCCAAAGGGAGTAAAGAGCTGATGAACATATGCA	1260
2698	CCGTAAGAAAAGGCTCTTAGAAGACTCCAAAGGGAGTAAAGAGCTGATGAACATATGCA	2757
1261	AGACCTCCAAGGTGAATTTGAAGCTCACAGAGTGTTTATCAAACTGGATGAAACAG	1320
2758	AGACCTCCAAGGTGAATTTGAAGCTCACAGAGTGTTTATCAAACTGGATGAAACAG	2817
1321	CCAAAAATCTCTGAGATCCCTGGAAAGGTTCCGATGATGCACTCTGTTACAAAGAGCTTT	1380
2818	CCAAAAATCTCTGAGATCCCTGGAAAGGTTCCGATGATGCACTCTGTTACAAAGAGCTTT	2877
1381	GGATAAATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCA	1440
2878	GGATAAATGAACCTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCA	2937
1441	TTTGGAAAGCAGTCTGACAGTGGAGAGCTGTCGACCTTTCTCTCAGAGAACTTCTGTT	1500
2938	TTTGGAAAGCAGTCTGACAGTGGAGAGCTGTCGACCTTTCTCTCAGAGAACTTCTGTT	2997
1501	GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTCC	1560
2998	GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTCC	3057
1561	AGCAGTTTCAAGACGACGATGTCATAGGCTTCAGAGGGAATGAAACTTAAGA	1620
3058	AGCAGTTTCAAGACGACGATGTCATAGGCTTCAGAGGGAATGAAACTTAAGA	3117
1621	ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACAGCCTTTGGA	1680
3118	ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACAGCCTTTGGA	3177
1681	AGACCTAGAGAAAATCTTACAGAGCCAGAGAGCTGCTCCTCAGAGAGAGAGCCCGAGAA	1740
3178	AGACCTAGAGAAAATCTTACAGAGCCAGAGAGCTGCTCCTCAGAGAGAGAGCCCGAGAA	3237
1741	TGTCACTCGGCTTCTACGAAACAGGCTGAGAGGTCAATTAAGTGGGAAATAATTGAA	1800
3238	TGTCACTCGGCTTCTACGAAACAGGCTGAGAGGTCAATTAAGTGGGAAATAATTGAA	3297
1801	CTTGCACTCGGCTGACTGACAGAGAAAATAGATGAGACCTCTGAAAGACTTCCAGGAAC	1860
3298	CTTGCACTCGGCTGACTGACAGAGAAAATAGATGAGACCTCTGAAAGACTTCCAGGAAC	3357
1861	TCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGGCCCAAGCTGAGGTATCAAGGATC	1920
3358	TCAAGAGGCCACGGATGAGCTGGAACCTCAAGCTGGCCCAAGCTGAGGTATCAAGGATC	3417
1921	CTGGCAGCCGCTGGGGGATCTCTCATTCAGTCTCTCAAGATCATCTCGAGAAAGTCAA	1980
3418	CTGGCAGCCGCTGGGGGATCTCTCATTCAGTCTCTCAAGATCATCTCGAGAAAGTCAA	3477
1981	GGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAACTGTAGGACCAAGTCAATGACCTTGC	2040
3478	GGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAACTGTAGGACCAAGTCAATGACCTTGC	3537

		LENGTH: 4848	
		TYPE: DNA	
		ORGANISM: Homo sapiens	
		US-09-845-416-35	
		Query Match 69.7%; Score 1443; DB 13; Length 4848;	
		Best Local Similarity 84.6%; Pred. No. 0;	
		Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;	
QY	1	GAGTATGCTACACACAGGCTGTTATGTCACACACCTCTGACCCCTACAGGAGCCCAT	60
DB	1680	GAGTATGCTACACACAGGCTGTTATGTCACACACCTCTGACCCCTACAGGAGCCCAT	1739
QY	61	TCCTTCACAGCATTTGGAAGCTCCTGAAGCAAGTCAATTTGGCAGTTCATTTGATGAGAG	120
DB	1740	TCCTTCACAGCATTTGGAAGCTCCTGAAGCAAGTCAATTTGGCAGTTCATTTGATGAGAG	1799
QY	121	TGAAGTAACTTGACCGTTTATCAACAGCTTTAGAGAGTATTTCGTGGCTTCTTTC	180
DB	1800	TGAAGTAACTTGACCGTTTATCAACAGCTTTAGAGAGTATTTCGTGGCTTCTTTC	1859
QY	181	TGCTGAGCACATTTGCAAGCACAGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA	240
DB	1860	TGCTGAGCACATTTGCAAGCACAGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA	1919
QY	241	CCAGTTTCATCTCATGAGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG	300
DB	1920	CCAGTTTCATCTCATGAGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG	1979
QY	301	TAATATTCTCAATTTGGAGTAGTACCTGATTGGAACAGGAAATTTATCAGAAGTGAAGA	360
DB	1980	TAATATTCTCAATTTGGAGTAGTACCTGATTGGAACAGGAAATTTATCAGAAGTGAAGA	2039
QY	361	AACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGAATGCTCAGGGTAGC	420
DB	2040	AACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGAATGCTCAGGGTAGC	2099
QY	421	TAGCATGGAATAAACAAGCATTTTACATAGATTTTAAATGATCTCCAGATCAGAACT	480
DB	2100	TAGCATGGAATAAACAAGCATTTTACATAGATTTTAAATGATCTCCAGATCAGAACT	2159
QY	481	GAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAAGAAACAGGAAATGAGGAAGA	540
DB	2160	GAAAGAGTTGAATGACTGGCTTAACAAAACAGAAAGAAAGAAACAGGAAATGAGGAAGA	2219
QY	541	GCCTCTTGACCTGATCTTGAAGACCTTAACAGCCCAAGTACACAAATAGGCTTCA	600
DB	2220	GCCTCTTGACCTGATCTTGAAGACCTTAACAGCCCAAGTACACAAATAGGCTTCA	2279
QY	601	AGAAGATCTAGAAACAGAAACAGGCTCAATTTCTCTCACTCAGATGCTGCTAGT	660
DB	2280	AGAAGATCTAGAAACAGAAACAGGCTCAATTTCTCTCACTCAGATGCTGCTAGT	2339
QY	661	TGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGGAAGAAACAATTAAGTATTGGG	720
DB	2340	TGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGGAAGAAACAATTAAGTATTGGG	2399
QY	721	AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAGACAT	780
DB	2400	AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAGAC--	2457
QY	781	CCTTCTCAAAATGGCAACGCTTTACTGAAAGACAGTGCCTTTTATGATGATGGCTTTCAGA	840
DB	2458	-----	2457
QY	841	AAAAGAGATGAGTGAACAAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAATGTT	900
DB	2458	-----	2457
QY	901	ATCAAGTCTTCAAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	960
DB	2458	-----	2457

QY	961	GGCAAACTGTATTCACTCAAAACAAGATCTTTCTTCAACACTGAAGATAAGTCAGTGAC	1020
DB	2458	-----	2457
QY	1021	CCAGAAGACGGAAGCATGGCTGGATTAATTTGGCCGGTGTGGGATAAATTTAGTCCAAA	1080
DB	2458	-----	2457
QY	1081	ACTTGAAGAAGAGTACAGCAGAGCTATAGATTATCTGCAACAGTTTCCCTTGACCTGGA	1140
DB	2458	-----ACTCATAGATTACTGCAACAGTTTCCCTTGACCTGGA	2495
QY	1141	AAAGTTTCTTCCCTGGCTTACAGAGCTGAAACAACCTGCAATGTCTTACAGATGCTAC	1200
DB	2496	AAAGTTTCTTCCCTGGCTTACAGAGCTGAAACAACCTGCAATGTCTTACAGATGCTAC	2555
QY	1201	CCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAAGAAACAATGGCA	1260
DB	2556	CCGTAAGGAAAGGCTCCTAGAGACTCCAAGGGAGTAAAGAGCTGATGAAGAAACAATGGCA	2615
QY	1261	AGACCTCCAAGGTGAATTTGAGCTCACAGATGTTTATCACAACTGATGAAGAAACAG	1320
DB	2616	AGACCTCCAAGGTGAATTTGAGCTCACAGATGTTTATCACAACTGATGAAGAAACAG	2675
QY	1321	CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTTCAAAAGACGTTT	1380
DB	2676	CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTTCAAAAGACGTTT	2735
QY	1381	GGATAACATGAATTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCTCA	1440
DB	2736	GGATAACATGAATTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCTCA	2795
QY	1441	TTTGAAGCCAGTTTCTGACAGTGGAGGCTCTGACCTTCTCTGAGGAACTTCTTGGT	1500
DB	2796	TTTGAAGCCAGTTTCTGACAGTGGAGGCTCTGACCTTCTCTGAGGAACTTCTTGGT	2855
QY	1501	GTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGGACCTTATGGAGCGACTTTCC	1560
DB	2856	GTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGGACCTTATGGAGCGACTTTCC	2915
QY	1561	AGCAGTTACAGAGCAGACGATGATAGGGCTTCAAGAGGGAATTTGAAAACTAAAGA	1620
DB	2916	AGCAGTTACAGAGCAGACGATGATAGGGCTTCAAGAGGGAATTTGAAAACTAAAGA	2975
QY	1621	ACCTGTAATCATGACTCTTTGAGACTGTGACATATTTCTGACAGAGAGCTTTTGA	1680
DB	2976	ACCTGTAATCATGACTCTTTGAGACTGTGACATATTTCTGACAGAGAGCTTTTGA	3035
QY	1681	AGGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCTCTGAGAGAGAGAGCCAGAA	1740
DB	3036	AGGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCTCTGAGAGAGAGAGCCAGAA	3095
QY	1741	TGTCAGTCTGGCTTTTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTGAA	1800
DB	3096	TGTCAGTCTGGCTTTTACGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAAATTGAA	3155
QY	1801	CCTGCACTCCGCTGATCTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACT	1860
DB	3156	CCTGCACTCCGCTGATCTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACT	3215
QY	1861	TCAAGAGGCAACGAGTGAAGTCAAGCTGCGCAAGCTGAGGCTGATCAAGGGATC	1920
DB	3216	TCAAGAGGCAACGAGTGAAGTCAAGCTGCGCAAGCTGAGGCTGATCAAGGGATC	3275
QY	1921	CTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAGATCACTCTGAGAGAACTCAA	1980
DB	3276	CTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAGATCACTCTGAGAGAACTCAA	3335
QY	1981	GGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCATGACTCTTGC	2040
DB	3336	GGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCATGACTCTTGC	3395
QY	2041	TCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCAACGGTATAACCTCAGCACTCTCGAAGA	2100

3396 TCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAGCTCTGCGAAGA 3455
2101 C 2101
3456 C 3456
RESULT 9
US-09-845-416-36
Sequence 36, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36
LENGTH: 5060
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-36
Query Match 68.7%; Score 1443; DB 13; Length 5060;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1777; Conservative 0; Mismatches 0; Indels 324; Gaps 1;
1 GAGCTATGCTCACACAGAGCTGCTTATGTGTCACACCTCTGACCCCTACACGAGGCCATT 60
1892 GAGCTATGCTCACACAGAGCTGCTTATGTGTCACACCTCTGACCCCTACACGAGGCCATT 1951
61 TCCTTACACAGCTTTGGAGCTCTCAAGCAAGTCAATTTGGCAGTTTCATTGATGAGAG 120
1952 TCCTTACACAGCTTTGGAGCTCTCAAGCAAGTCAATTTGGCAGTTTCATTGATGAGAG 2011
121 TGAAGTAAACCTGGACCGTTTATCAACACAGCTTTAGAGAGTATTATCGTGGCTCTTTTC 180
2012 TGAAGTAAACCTGGACCGTTTATCAACACAGCTTTAGAGAGTATTATCGTGGCTCTTTTC 2071
181 TGTGAGGACACATTTGAAGCAACAGAGAGATTTCTTAATGATGGAAGTGGTGAAGA 240
2072 TGTGAGGACACATTTGAAGCAACAGAGAGATTTCTTAATGATGGAAGTGGTGAAGA 2131
241 CCAGTTTCACTACTCATGAGGGGTACATGATGATTTTGAAGCCCATCAGGCGCGGTTGG 300
2132 CCAGTTTCACTACTCATGAGGGGTACATGATGATTTTGAAGCCCATCAGGCGCGGTTGG 2191
301 TAATATTCTCAATTTGGGAAGTAAAGTGTGTAAGCAAGGAAATTTATCAGAGATGAAGA 360
2192 TAATATTCTCAATTTGGGAAGTAAAGTGTGTAAGCAAGGAAATTTATCAGAGATGAAGA 2251
361 AACTGAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGGATGCTCAGGAGTAC 420
2252 AACTGAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGGATGCTCAGGAGTAC 2311
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTGTTTAATGGATCTCCAGAACTCAGAACT 480
2312 TAGCATGGAAAAACAAAGCAATTTACATAGAGTGTTTAATGGATCTCCAGAACTCAGAACT 2371
481 GAAAGAGTGAATGACTGGCTTACAAAAACAGAGAAAGAACAGGAAATTTGAGGAGGAAGA 540
2372 GAAAGAGTGAATGACTGGCTTACAAAAACAGAGAAAGAACAGGAAATTTGAGGAGGAAGA 2431
541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAGTACAAACATTAAGTGTCTTCA 600
2432 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAGTACAAACATTAAGTGTCTTCA 2491

QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACTCACTGGTGGTGTAGT 560
DB 2492 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACTCACTGGTGGTGTAGT 2551
QY 661 TGAATGATCTAGTGGAGATCAGCAACTCTGCTTTGGAGAACCACTTTAAGGTATTTGG 720
DB 2552 TGAATGATCTAGTGGAGATCAGCAACTCTGCTTTGGAGAACCACTTTAAGGTATTTGG 2611
QY 721 AGATCGATGGCAAAACATCTGTAGATGGAACAAGACCGCTGGTTCCTTTTACAAGACAT 780
DB 2612 AGATCGATGGCAAAACATCTGTAGATGGAACAAGACCGCTGGTTCCTTTTACAAGAC-- 2669
QY 781 CCTTCTCAATGSCAACGCTTCTACTGGAAGNACAGTGCCTTTTGTAGTGCATGGCTTTTCA 840
DB 2670 ----- 2669
QY 841 AAAAGAAGATGCAGTGAACAAGATTACACAACCTGGCTTTTAAAGATCAAAATGATTT 900
DB 2670 ----- 2669
QY 901 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
DB 2670 ----- 2669
QY 961 GGGCAAACTGTATTCTCACTCAAAACAGATCTTCTTCAACACTGAAGATAAAGTCAGTGAC 1020
DB 2670 ----- 2669
QY 1021 CCAGAAGACGGAAGAGCTGGCTGGATTAACCTTTGCCCGGTGTGGGNTAATTTAGTCCAAA 1080
DB 2670 ----- 2669
QY 1081 ACTTGAAGAGGTACAGCAGACAGACTCATAGATTACTGCAACAGTTCCTCCCTGACCTGGGA 1140
DB 2670 -----ACTCATAGATTACTGCAACAGTTCCTCCCTGACCTGGGA 2707
QY 1141 AAAGTTTCTTGGCTGTACAGAAGCTGAAAACAATGCCAATGTCTCTACAGGATGCTAC 1200
DB 2708 AAAGTTTCTTGGCTGTACAGAAGCTGAAAACAATGCCAATGTCTCTACAGGATGCTAC 2767
QY 1201 CCCTGAAGAAAGCTCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCA 1260
DB 2768 CCCTGAAGAAAGCTCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCA 2827
QY 1261 AGACCTCCAAAGGTGAAATTTGAAGCTCAACAGATGTTTATCACAACCTGGATGAAACAG 1320
DB 2828 AGACCTCCAAAGGTGAAATTTGAAGCTCAACAGATGTTTATCACAACCTGGATGAAACAG 2887
QY 1321 CCAAAAATCTTGAGTCCCTGGAAGSTTCGGATGATGAGTCTCTGTACAGAGACGTTT 1380
DB 2888 CCAAAAATCTTGAGTCCCTGGAAGSTTCGGATGATGAGTCTCTGTACAGAGACGTTT 2947
QY 1381 GGATAACATGAACCTTCAAGTGGAGTGAACCTTGGAAAAAGTCTCTCAACATTAAGTCCCA 1440
DB 2948 GGATAACATGAACCTTCAAGTGGAGTGAACCTTGGAAAAAGTCTCTCAACATTAAGTCCCA 3007
QY 1441 TTTGGAAGCCAGTTTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAAGGAACTTCTGCT 1500
DB 3008 TTTGGAAGCCAGTTTCTGACCAAGTGAAGCGTCTGCACTTTCTCTGCAAGGAACTTCTGCT 3067
QY 1501 GTGGCTACAGCTGAAGATGATGAAATTTAAGCCGCGAGCACCTATTGGAGGCGACTTTCC 1560
DB 3068 GTGGCTACAGCTGAAGATGATGAAATTTAAGCCGCGAGCACCTATTGGAGGCGACTTTCC 3127
QY 1561 AGCAGTTTCAAGAGCAGAACCGATGTACATAGGGCTTCAAGAGGGAATTTGAAAATAAGA 1620
DB 3128 AGCAGTTTCAAGAGCAGAACCGATGTACATAGGGCTTCAAGAGGGAATTTGAAAATAAGA 3187
QY 1621 ACCTGTATCATGTAGTACTCTTGAGACTGTAGCAATATTTCTGACAGAGCAGCTTTTGA 1680
DB 3188 ACCTGTATCATGTAGTACTCTTGAGACTGTAGCAATATTTCTGACAGAGCAGCTTTTGA 3247
QY 1681 AGGACTAGAGAAACTCTACCAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGAGAGAGAG 1740

3248 AGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTGAGGAGAGAGCCAGAA 3307
1741 TGTCACTCGCTTCTACGAAGCAGGCTGAGAGGCTCAATCTGAGTGGGAAATTGAA 1800
3308 TGTCACTCGCTTCTACGAAGCAGGCTGAGAGGCTCAATCTGAGTGGGAAATTGAA 3367
1801 CTGCACTCGCTGCTGAGTGGCAGAGAAATAGATGAGACCCCTTGAAGACTCCAGAACT 1860
3368 CTGCACTCGCTGCTGAGTGGCAGAGAAATAGATGAGACCCCTTGAAGACTCCAGAACT 3427
1861 TCAGAGGCCACGGATGAGCTGAGCCTCAAGCTGCGCAGCTGAGTGAAGGATC 1920
3428 TCAAGAGGCCACGGATGAGCTGAGCCTCAAGCTGCGCAGCTGAGTGAAGGATC 3487
1921 CTGCACTCGCTGCTGAGTGGCAGTCTCTCAATGACCTCTCTCAAGATCACTCGAGAAAGTCAA 1980
3488 CTGCACTCGCTGCTGAGTGGCAGTCTCTCAATGACCTCTCTCAAGATCACTCGAGAAAGTCAA 3547
1981 GGCACCTCGAGGAAATTTGGGCTCTGAAAGAGACGTCGAGCAGCTCAATGACCTTGC 2040
3548 GGCACCTCGAGGAAATTTGGGCTCTGAAAGAGACGTCGAGCAGCTCAATGACCTTGC 3607
2041 TCGCCAGCTTACCACTTTGGGCTCTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
3608 TCGCCAGCTTACCACTTTGGGCTCTGAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3667
2101 C 2101
3668 C 3668

RESULT 10
US-10-149-736-42
; Sequence 42, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149, 736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238, 848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-42
Query Match 61.5%; Score 1293; DB 13; Length 8689;
Best Local Similarity 72.5%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 798; Gaps 1;
QY 1 GAGCTATGCTTACACAGAGCTGCTTATGTCACACCTCTGACCCCTACAGGAGCCCAT 60
DB 1099 GAGCTATGCTTACACAGAGCTGCTTATGTCACACCTCTGACCCCTACAGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCTGAGACAAAGTCAATTTGGCAGTTCATTGATGAGAG 120
DB 1159 TCCTTCACAGCATTTGGAAGCTCTGAGACAAAGTCAATTTGGCAGTTCATTGATGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAGAGAGTATTATCGTGGCTTTTC 180
DB 1219 TGAAGTAAACCTGGACCGCTTATCAAAACAGCTTTAGAGAGAGTATTATCGTGGCTTTTC 1278

QY 181 TGTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
DB 1279 TGTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
QY 241 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 300
DB 1339 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 1398
QY 301 TAATAATTCTACAATTCGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 360
DB 1399 TAATAATTCTACAATTCGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 1458
QY 361 AACTGAGTACAGAGACAGATGATCTCTAATTTCAAGATGGATGCTCTCAGGCTAGC 420
DB 1459 AACTGAGTACAGAGACAGATGATCTCTAATTTCAAGATGGATGCTCTCAGGCTAGC 1518
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAACT 480
DB 1519 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAACAAAGGAAATGAGGAAAGA 540
DB 1579 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAACAAAGGAAATGAGGAAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACATAAGGTGCTTCA 600
DB 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACATAAGGTGCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGACAGTCAAGGTCAATTTCTCTCCTCAGATGGTGGTGTAGT 660
DB 1699 AGAAGATCTAGAACAAAGACAGTCAAGGTCAATTTCTCTCCTCAGATGGTGGTGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAAAGAACAACTTAAAGTATTGGG 720
DB 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAAAGAACAACTTAAAGTATTGGG 1818
QY 721 AGATCGATGGGCAAAACATCTGTAGATGGAAGAACCGCTGGGTCTTTTACAGACAT 780
DB 1819 AGATCGATGGGCAAAACATCTGTAGATGGAAGAACCGCTGGGTCTTTTACAGACAT 1878
QY 781 CCTTCTCAATGCGCAAGCTCTTCTGAGAACAGTGCCTTTTGTGATGCTGCTTTTACAGA 840
DB 1879 CCTTCTCAATGCGCAAGCTCTTCTGAGAACAGTGCCTTTTGTGATGCTGCTTTTACAGA 1938
QY 841 AAAAGAGATGCAAGTGAACAAAGATTCACAACTGGCTTTAAAGATCAAAATGAATGTT 900
DB 1939 AAAAGAGATGCAAGTGAACAAAGATTCACAACTGGCTTTAAAGATCAAAATGAATGTT 1998
QY 901 ATCAAGTCTTCAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAGAAAAAGCAATCCAT 960
DB 1999 ATCAAGTCTTCAAAACCTGGCGCTTTTAAAGCGGATCTAGAAAGAAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCTCAAAACAGATCTTTTCAACACTGAAGAATAAGTCAAGTAC 1020
DB 2059 GGGCAAACTGTATTCTCAAAACAGATCTTTTCAACACTGAAGAATAAGTCAAGTAC 2118
QY 1021 CCAGAGACGGAAGCATGCTGATTAATTTGCGCGGTGTTGGGATAATTTAGTCCAAA 1080
DB 2119 CCAGAGACGGAAGCATGCTGATTAATTTGCGCGGTGTTGGGATAATTTAGTCCAAA 2178
QY 1081 ACTTGAAGAGAGTACAGACAG----- 1102
DB 2179 ACTTGAAGAGAGTACAGACAGATTTTACAGAGCCCTGACCTAGTCTCTGACTGACCAC 2238
QY 1103 ----- 1102
DB 2239 TATTGGAGCTCTCTCTACTCAGACTGTTACTTGTGTGACAAACCTGTGTGTTACTAAGGA 2298
QY 1103 ----- 1102
DB 2299 AACTGCCATCTCCAAACTAGAAATGCCATCTTCTTGTGATGTTGGAGTACCTGCTCTGGC 2358
QY 1103 ----- 1102

Db 2359 AGATTTCACCCGGCTTGGACAGAACTTACCGACTGGCTTCTCTGCTGTATCAAGTTAT 2418
QY 1103 ----- 1102
Db 2419 AAAATCACAGAGGGTGATGGTGGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCA 2478
QY 1103 ----- 1102
Db 2479 GAAGCAACAATGCAGGATTTGGAACACAGAGCGTCCCGAGTTGGAAGAACTCATTTACCGC 2538
QY 1103 ----- 1102
Db 2539 TGCCCAAAATTTGAARACAAAGACCAGCAATCAAGAGGCTAGAACAAATCATTTACGGATCG 2598
QY 1103 ----- 1102
Db 2599 AATTGAAAGAAATTCAGAATCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGAGCA 2658
QY 1103 ----- 1102
Db 2659 ACAGTTGATGAATGTTAAAGGATTCACAACAATGGCTGGAAGCTTAAGGAAGAGCTGA 2718
QY 1103 ----- 1102
Db 2719 GCAGGTCTTAGGACAGGCCAGACCAAGCTTGAGTCATGGAAGGAGGTCCCTATACAGT 2778
QY 1103 ----- 1102
Db 2779 AGATGCAATCCAAAAGAAATTCAGAAACCAAGCAGTTGGCCAAAGACCTCGGCCAGTG 2838
QY 1103 ----- 1102
Db 2839 GCAGCAAAATGTAGATGTGGCAATGACTTGGCCCTGAAACTTCTCCGGATTAATTCTGC 2898
QY 1103 ----- 1102
Db 2899 AGATGATACCAGAAAAGTCCACATGATATAACAGAGAATATCAATGCCCTCTTGGAGAAGCAT 2958
QY 1103 ----- 1102
Db 2959 TCATPAAAGGTGAGTACGAGGAGGCTGCTTTGGAAGAACTCATAGATTCTGCAACA 3018
QY 1123 GTTCCCTCGACTGGAAGAGTTTCTTGCTGGCTTACAGAGCTGAAACAACTCCCAA 1182
Db 3019 GTTCCCTCGACTGGAAGAGTTTCTTGCTGGCTTACAGAGCTGAAACAACTCCCAA 3078
QY 1183 TGTCTACAGATCTACCGTAAAGGAGGCTCCTAGAGACTCCAGGAGTAAAGA 1242
Db 3079 TGTCTACAGATCTACCGTAAAGGAGGCTCCTAGAGACTCCAGGAGTAAAGA 3138
QY 1243 GCTGATGAAACAAATGGCAAGACTCCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCA 1302
Db 3139 GCTGATGAAACAAATGGCAAGACTCCCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCA 3198
QY 1303 CAACCTGGATGAAACAGCAACAAATTCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 1362
Db 3199 CAACCTGGATGAAACAGCAACAAATTCCTGAGATCCCTGGAAGGTTCCGATGATGCAGT 3258
QY 1363 CCTGTTACAAAGAGCTTTGGATACATGAATCAAGTGGAGTGAATTCGGAAGAAAGTC 1422
Db 3259 CCTGTTACAAAGAGCTTTGGATACATGAATCAAGTGGAGTGAATTCGGAAGAAAGTC 3318
QY 1423 TCTCAACATTAGTCCCATTTGGAGCCAGTCTTGACAGTGAAGCGTGTGACGCTTTC 1482
Db 3319 TCTCAACATTAGTCCCATTTGGAGCCAGTCTTGACAGTGAAGCGTGTGACGCTTTC 3378
QY 1483 TCTCAGCAACTCTTGCTGGCTTACAGCTGAAGATGATGAATTAAGCGGAGGCAACC 1542
Db 3379 TCTCAGCAACTCTTGCTGGCTTACAGCTGAAGATGATGAATTAAGCGGAGGCAACC 3438
QY 1543 TATTGGAGGCGACTTTCAGCAGTTTACAGAGCAGAACGATGTACATAGGCGCTTCAAGAG 1602

Db 3439 TATTGGAGGCGACTTTCAGCAGTTCCAGCAGTTCAAGACAGAACGATGTACATAGGCGCTTCAAGAG 3498
QY 1603 GGAATTTGAAACTTAAAGAACCTGTAAATCATGAGTACTCTTGAGAGCTGTACGAATATTTCT 1662
Db 3499 GGAATTTGAAACTTAAAGAACCTGTAAATCATGAGTACTCTTGAGAGCTGTACGAATATTTCT 3558
QY 1663 GACAGAGAGCCCTTTTGGAAAGGACTAGAGAACTCTTACAGGAGCCAGAGAGCTGCCTCC 1722
Db 3559 GACAGAGAGCCCTTTTGGAAAGGACTAGAGAACTCTTACAGGAGCCAGAGAGCTGCCTCC 3618
QY 1723 TGAGAGAGAGCCCAAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATAC 1782
Db 3619 TGAGAGAGAGCCCAAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATAC 3678
QY 1783 TGAGTGGGAAAAAATTTGAACCTTCGACTCCGCTGAGTGGCAGAGAAAAATAGATGAGACCT 1842
Db 3679 TGAGTGGGAAAAAATTTGAACCTTCGACTCCGCTGAGTGGCAGAGAAAAATAGATGAGACCT 3738
QY 1843 TGAAAGACTTCCAGGAACTTCAAGAGCCACGATGAGTGGACCTCAAGCTGCCTCAAGC 1902
Db 3739 TGAAAGACTTCCAGGAACTTCAAGAGCCACGATGAGTGGACCTCAAGCTGCCTCAAGC 3798
QY 1903 TGAGTGTATCAAGGAGTCTTGGCAGCCGCTGGGCGATCTCCTCATTTGACTCTCTCCAAGA 1962
Db 3799 TGAGTGTATCAAGGAGTCTTGGCAGCCGCTGGGCGATCTCCTCATTTGACTCTCTCCAAGA 3858
QY 1963 TCACCTCAGAAAAGTCAAGGACCTTCGAGGAGAAATTTGGGCTCTGAAAGAGACGTGAG 2022
Db 3859 TCACCTCAGAAAAGTCAAGGACCTTCGAGGAGAAATTTGGGCTCTGAAAGAGACGTGAG 3918
QY 2023 CCACGCTCAATGACCTTGTCTCGCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAA 2082
Db 3919 CCACGCTCAATGACCTTGTCTCGCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAA 3978
QY 2083 CCTCAGCACTCTGGAAGAC 2101
Db 3979 CCTCAGCACTCTGGAAGAC 3997

RESULT 11
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR FILING DATE: 60/200,777
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match 53.0%; Score 1114; DB 13; Length 4414;
Best Local Similarity 97.0%; Pred. No. 2.3e-313;
Matches 1135; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 GAGCTATCCCTACACAGAGGCTGTATGTACCAACCTCTGACCTACAGGAGCCCAT 60
Db 1657 GAGCTATCCCTACACAGAGGCTGTATGTACCAACCTCTGACCTACAGGAGCCCAT 1716
QY 61 TCCTTCACAGCAATTTGGAAGCTCCTGGAAGCAAGTCAATTTGGCAGTTCATTTGAGAG 120
Db 1717 TCCTTCACAGCAATTTGGAAGCTCCTGGAAGCAAGTCAATTTGGCAGTTCATTTGAGAG 1776
QY 121 TGAAGTAAACCTGAGCCGTTATCAACAGCTTTAGAGAAGATATTATCGTGGCTTCTTC 180

1777 TGAAGTAACTGACCGTATTCAACAGCTTTAGAAAGATATTATCGTGGCTTCTTC 1836
181 TGCTGAGGACATTTGACGACACAGGAGAGATTTCTTAATGATGTGGAAGTGAAGA 240
1837 TGCTGAGGACATTTGACGACACAGGAGAGATTTCTTAATGATGTGGAAGTGAAGA 1896
241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 300
1897 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 1956
301 TAAATTTCTCAATTTGGAGTAACTGATTTGACAGGAAATTTATCAGAAATGAAGA 360
1957 TAAATTTCTCAATTTGGAGTAACTGATTTGACAGGAAATTTATCAGAAATGAAGA 2016
361 AACTGAAGTCAAGAGCAGATGAATCTCTAAATTTCAAGATGGAATGCCCTCAGGGTAGC 420
2017 AACTGAAGTCAAGAGCAGATGAATCTCTAAATTTCAAGATGGAATGCCCTCAGGGTAGC 2076
421 TAGCATGGAATAAACAAGCAATTTACATAGATTTTAAATGATCTCCAGAACT 480
2077 TAGCATGGAATAAACAAGCAATTTACATAGATTTTAAATGATCTCCAGAACT 2136
481 GAAAGATTGAATGACTGGCTTAAACAAACAGAAAGAAAGCAAGGAAATGACGAAGA 540
2137 GAAAGATTGAATGACTGGCTTAAACAAACAGAAAGAAAGCAAGGAAATGACGAAGA 2196
541 GCCTTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACAAACAATAGGTGCTTCA 600
2197 GCCTTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACAAACAATAGGTGCTTCA 2256
601 AGAAGATTGAATGACTGGCTTAAACGCAAGTACAAACAATAGGTGCTTCA 660
2257 AGAAGATTGAATGACTGGCTTAAACGCAAGTACAAACAATAGGTGCTTCA 2316
661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAAACAATTAAGTATTGGG 720
2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAAACAATTAAGTATTGGG 2376
721 AGATCGATGGCAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTACAGACAT 780
2377 AGATCGATGGCAACATCTGTAGATGGACAGAGACCGCTGGTCTTTTACAGACAT 2436
781 CCTTCTCAAAATGGCAACCTTCTACTGAAGAACAGTGCCTTTTGTGCTGCTTTTCA 840
2437 CCTTCTCAAAATGGCAACCTTCTACTGAAGAACAGTGCCTTTTGTGCTGCTTTTCA 2496
841 AAAAGAGATGCGATGAACAGATTTCAACAATGCTTTTAAAGATCAAAATGAATGTT 900
2497 AAAAGAGATGCGATGAACAGATTTCAACAATGCTTTTAAAGATCAAAATGAATGTT 2556
901 ATCAAGTCTTCAAAACCTGGCCGTTTAAAGCGGATCTAGAAAGAAAGCAATCCAT 960
2557 ATCAAGTCTTCAAAACCTGGCCGTTTAAAGCGGATCTAGAAAGAAAGCAATCCAT 2616
961 GGGCAAACTGATTTCACTCAAAACAGATCTTCTTTCAACACTGAAGATAAGTCACTGAC 1020
2617 GGGCAAACTGATTTCACTCAAAACAGATCTTCTTTCAACACTGAAGATAAGTCACTGAC 2676
1021 CCAGAAGACGGAACGATGGCTGATTAATTTGCGCGGTGTTGGATTAATTTAGTCCAAA 1080
2677 CCAGAAGACGGAACGATGGCTGATTAATTTGCGCGGTGTTGGATTAATTTAGTCCAAA 2736
1081 ACTTGAAGAGTACAGCAGACTCATAGATTTACTGCAACAGTCTCCCTCCCTGACCTGGA 1140
2737 ACTTGAAGAGTACAGCAGACTCATAGATTTACTGCAACAGTCTCCCTCCCTGACCTGGA 2796
1141 AAGTTTCTTGGCTTACAGAGCTGA 1170
2797 TGAGCTGACCTCAAGCTGGCCCAAGCTGA 2826

RESULT 12

US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 5417
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-39
Query Match 52.5%; Score 1103.4; DB 13; Length 5417;
Best Local Similarity 99.9%; Pred. No. 3.3e-310; Mismatches 1; Indels 0; Gaps 0;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATCCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 60
DB 1099 GAGCTATCCCTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCCTGGAAGCAAGTCAATTTGGGAGTTCATTGATGGAGAG 120
DB 1159 TCCTTCACAGCATTTGGAAGCTCCTGGAAGCAAGTCAATTTGGGAGTTCATTGATGGAGAG 1218
QY 121 TGAAGTAACTCGACCGTATCAACACAGCTTTAGAAAGAGTATTTCGTGGCTTCTTTC 180
DB 1219 TGAAGTAACTCGACCGTATCAACACAGCTTTAGAAAGAGTATTTCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACATTTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGAAGA 240
DB 1279 TGCTGAGGACATTTGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGAAGA 1338
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGCGCGGTGG 300
DB 1339 CCAGTTTCATCTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGCGCGGTGG 1398
QY 301 TAAATTTCTCAAAATGGGAAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAATGAAGA 360
DB 1399 TAAATTTCTCAAAATGGGAAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAATGAAGA 1458
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCCTCAGGTAGC 420
DB 1459 AACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCCTCAGGTAGC 1518
QY 421 TAGCATGGAATAAACAAGCAATTTACATAGATTTTAAATGATCTCCAGAACTCAGAACT 480
DB 1519 TAGCATGGAATAAACAAGCAATTTACATAGATTTTAAATGATCTCCAGAACTCAGAACT 1578
QY 481 GAAAGAGTGAATGACTGGCTTAAACAAACAGAAAGAAAGCAAGGAAATGAGGAGAGA 540
DB 1579 GAAAGAGTGAATGACTGGCTTAAACAAACAGAAAGAAAGCAAGGAAATGAGGAGAGA 1638
QY 541 GCCTTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACAAACAATAGGTGCTTCA 600
DB 1639 GCCTTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACAAACAATAGGTGCTTCA 1698
QY 601 AGAAGATTGAAGCAAGCAAGTCAAGGTCAATTTCTCACTCAATGTTGGTGTAGT 660
DB 1699 AGAAGATTGAAGCAAGCAAGTCAAGGTCAATTTCTCACTCAATGTTGGTGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAAGAAACAATTAAGGTATTGGG 720

Db	1759	TGATGATCTAGTGGAGATCACCAACTGCTGTTTGGAGAACAACTTAAGGTATGGG	1818
Qy	721	AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT	780
Db	1819	AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT	1878
Qy	781	CTTCTCAAAATGCGCAAGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCAGCTTTTACA	840
Db	1879	CTTCTCAAAATGCGCAAGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCAGCTTTTACA	1938
Qy	841	AAAGAGATGTCAGTGAACAGATTCACACAACTGGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	900
Db	1939	AAAGAGATGTCAGTGAACAGATTCACACAACTGGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	1998
Qy	901	ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	960
Db	1999	ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	2058
Qy	961	GGGCAAACTGTATCTACTCAAAAGATCTTCTTCAACACTGACAGAAATAAGTCAGTGAC	1020
Db	2059	GGGCAAACTGTATCTACTCAAAAGATCTTCTTCAACACTGACAGAAATAAGTCAGTGAC	2118
Qy	1021	CCAGAGACGGAAGCATGGCTGGATAACTTTGGCCGGTGTGGGATAATTTAGTCCAAA	1080
Db	2119	CCAGAGACGGAAGCATGGCTGGATAACTTTGGCCGGTGTGGGATAATTTAGTCCAAA	2178
Qy	1081	ACTTGAAGAGATGACAGACAGACT	1105
Db	2179	ACTTGAAGAGATGACAGACAGACT	2203
RESULT 13			
US-10-149-736-44			
; Sequence 44, Application US/10149736			
; Publication No. US20030216332A1			
; GENERAL INFORMATION:			
; APPLICANT: Chamberlain, Jeffrey S.			
; APPLICANT: Harper, Scott Q.			
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences			
; FILE REFERENCE: UN-06968			
; CURRENT APPLICATION NUMBER: US/10/149,736			
; CURRENT FILING DATE: 2002-06-17			
; PRIOR APPLICATION NUMBER: PCT/US01/31126			
; PRIOR FILING DATE: 2001-10-04			
; PRIOR APPLICATION NUMBER: 60/238,848			
; PRIOR FILING DATE: 2000-10-06			
; NUMBER OF SEQ ID NOS: 96			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 44			
; LENGTH: 11443			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Synthetic			
US-10-149-736-44			
Query Match			
Best Local Similarity 99.9%; Score 1103.4; DB 13; Length 11443;			
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	GAGTATGCTTACACACAGCGTGTATGTACACACCTCTGACCCCTACACGAGCCCAT	60
Db	1099	GAGTATGCTTACACACAGCGTGTATGTACACACCTCTGACCCCTACACGAGCCCAT	1158
Qy	61	TCCTTCACAGATTTGGAAGCTCTGAGACAGTCAATTTGGCAGTTCATTTGATGAGAG	120
Db	1159	TCCTTCACAGATTTGGAAGCTCTGAGACAGTCAATTTGGCAGTTCATTTGATGAGAG	1218
Qy	121	TGAAGTAAACCTGGACCGTGTATCAACACAGCTTTTAGAAGAGATTTATCGTGGCTTTCTT	180
Db	1219	TGAAGTAAACCTGGACCGTGTATCAACACAGCTTTTAGAAGAGATTTATCGTGGCTTTCTT	1278

Qy	181	TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTATGATGTGGAGTGGTGAAGA	240
Db	1279	TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTATGATGTGGAGTGGTGAAGA	1138
Qy	241	CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG	300
Db	1339	CCAGTTTCATCTACTCATGAGGGGTACATGATGGATTTTGACAGCCCATCAGGGCCGGGTGG	1398
Qy	301	TAATATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGAAAAATTTACAGAGATGA	360
Db	1399	TAATATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGAAAAATTTACAGAGATGA	1458
Qy	361	AACCTCAAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGATGCTCAGGGTAGC	420
Db	1459	AACCTCAAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGATGCTCAGGGTAGC	1518
Qy	421	TAGCATGGAAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACT	480
Db	1519	TAGCATGGAAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACT	1578
Qy	481	GAAGAGTTGATGATCTGCTTAAAGAACAGAAAAACAAGAAAAATGAGGAGAGA	540
Db	1579	GAAGAGTTGATGATCTGCTTAAAGAACAGAAAAACAAGAAAAATGAGGAGAGA	1638
Qy	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAGTACAAACAACATAGGTGCTTCA	600
Db	1639	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCAGTACAAACAACATAGGTGCTTCA	1698
Qy	601	AGAAGATCTAGAACAAACAAGTCAGGGTCAATTTCTCTCACTCACATGTTGGTGGT	660
Db	1699	AGAAGATCTAGAACAAACAAGTCAGGGTCAATTTCTCTCACTCACATGTTGGTGGT	1758
Qy	661	TGATGAATCTAGTGGAGATCACCGCACTGCTTTTGGAGAACCACTTAAGGTATTGGG	720
Db	1759	TGATGAATCTAGTGGAGATCACCGCACTGCTTTTGGAGAACCACTTAAGGTATTGGG	1818
Qy	721	AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT	780
Db	1819	AGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT	1878
Qy	781	CCTTCTCAATGGCAACGCTCTTACTGAAGACAGTGCCTTTTGTAGTCATGGCTTTCAGA	840
Db	1879	CCTTCTCAATGGCAACGCTCTTACTGAAGACAGTGCCTTTTGTAGTCATGGCTTTCAGA	1938
Qy	841	AAAAGAGATGCTAGTGAACCAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAATGTT	900
Db	1939	AAAAGAGATGCTAGTGAACCAAGATTTCACAACTGGCTTTTAAAGATCAAAATGAATGTT	1998
Qy	901	ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	960
Db	1999	ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	2058
Qy	961	GGGCAAACTGTATTCACCTCAAAACAGATCTTTTCAACACTGAAGATAAGTCAGTGAC	1020
Db	2059	GGGCAAACTGTATTCACCTCAAAACAGATCTTTTCAACACTGAAGATAAGTCAGTGAC	2118
Qy	1021	CCAGAGACGGAAGCATGGCTGGATAACTTTGGCCGGTGTGGGATAATTTAGTCCAAA	1080
Db	2119	CCAGAGACGGAAGCATGGCTGGATAACTTTGGCCGGTGTGGGATAATTTAGTCCAAA	2178
Qy	1081	ACTTGAAGAGTACAGACAGACT	1105
Db	2179	ACTTGAAGAGTACAGACAGACT	2203
RESULT 14			
US-10-149-736-47			
; Sequence 47, Application US/10149736			
; Publication No. US20030216332A1			
; GENERAL INFORMATION:			
; APPLICANT: Chamberlain, Jeffrey S.			
; APPLICANT: Harper, Scott Q.			
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences			

FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1
SEQ ID NO 47
LENGTH: 12057
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-149-736-47

Query Match 52.5%; Score 1103.4; DB 13; Length 12057;
Best Local Similarity 99.9%; Pred. No. 5.6e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCTACACACAGCGTCTTATGTCTACACACCTCTGACCCCTACACGAGCCCAT 60
DB 1099 GAGCTATGCTACACACAGCGTCTTATGTCTACACCTCTGACCCCTACACGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
DB 1159 TCCTTCACAGCATTTGGAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 1218
QY 121 TGAAGTAACTGGACCGTTTCAACAGCTTTAGAACAGTATATCGTGGCTCTTTTC 180
DB 1219 TGAAGTAACTGGACCGTTTCAACAGCTTTAGAACAGTATATCGTGGCTCTTTTC 1278
QY 181 TCCTGAGGACACATTTGGAAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 240
DB 1279 TCCTGAGGACACATTTGGAAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 1338
QY 241 CCAGTTTCACTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGCGCGGTTGG 300
DB 1339 CCAGTTTCACTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGCGCGGTTGG 1398
QY 301 TAATATTTCACTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGCGCGGTTGG 360
DB 1399 TAATATTTCACTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGCGCGGTTGG 1458
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTAAATCAAGATGGGAATCCCTCAGGTTAGC 420
DB 1459 AACTGAAGTACAGAGCAGATGAATCTCTAAATCAAGATGGGAATCCCTCAGGTTAGC 1518
QY 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 480
DB 1519 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAACAGAGAAAGAAAGAAAGAAAGTGGAGGA 540
DB 1579 GAAAGAGTTGAATGACTGGCTTAACAAAACAGAGAAAGAAAGAAAGTGGAGGA 1638
QY 541 GCCTCTGACAGCTGATCTTGAAGACCTAAACCGCAAGTACAGCACTAAGAGTCTTCA 600
DB 1639 GCCTCTGACAGCTGATCTTGAAGACCTAAACCGCAAGTACAGCACTAAGAGTCTTCA 1698
QY 601 AGAAGATCTAGAACAGAAACAAGTCAAGGTCAATTTCTCTCACTCAATGGTGGTGTAGT 660
DB 1699 AGAAGATCTAGAACAGAAACAAGTCAAGGTCAATTTCTCTCACTCAATGGTGGTGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCAAGCACTGCTCTTTGGAGAAAGAACTTAAGGTATTGGG 720
DB 1759 TGATGAATCTAGTGGAGATCAAGCACTGCTCTTTGGAGAAAGAACTTAAGGTATTGGG 1818
QY 721 AGATTCATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAGACAT 780
DB 1819 AGATTCATGGGCAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAGACAT 1878

QY 781 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCA 840
DB 1879 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCA 1938
QY 841 AAAAGAAGATGAGTGAACAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
DB 1939 AAAAGAAGATGAGTGAACAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
DB 1999 ATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCCTCAAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 1020
DB 2059 GGGCAAACTGTATTCCTCAAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGAC 2118
QY 1021 CCAGAAGAGCGAAGCATGGCTGGATACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
DB 2119 CCAGAAGAGCGAAGCATGGCTGGATACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 2178
QY 1081 ACTTGAAGAGGTACAGCACAGACT 1105
DB 2179 ACTTGAAGAGGTACAGCACAGACT 2203
RESULT 15
US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 13957
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-22
Query Match 52.5%; Score 1103.4; DB 10; Length 13957;
Best Local Similarity 99.9%; Pred. No. 6.2e-310;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCTACACACAGCGTCTTATGTCTACACCTCTGACCCCTACACGAGCCCAT 60
DB 1099 GAGCTATGCTACACACAGCGTCTTATGTCTACACCTCTGACCCCTACACGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 120
DB 1159 TCCTTCACAGCATTTGGAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 1218
QY 121 TGAAGTAACTGGACCGTTTCAACAGCTTTAGAACAGTATATCGTGGCTCTTTTC 180
DB 1219 TGAAGTAACTGGACCGTTTCAACAGCTTTAGAACAGTATATCGTGGCTCTTTTC 1278
QY 181 TCCTGAGGACACATTTGGAAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 240
DB 1279 TCCTGAGGACACATTTGGAAGCTCCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGGAGAG 1338
QY 241 CCAGTTTCACTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGCGCGGTTGG 300
DB 1339 CCAGTTTCACTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGCGCGGTTGG 1398
QY 301 TAATATTTCACTCATGAGGGGTACATGATGGATTTCAGAGCCCATCAGGCGCGGTTGG 360

Db	1399	TAATATTCTACAATTGGGAAGTAAGCTGATTGGAAACAGGAAAATATACAGAGATGAAGA	1458
2y	361	AACCTGAAGTACAGAGACAGATGAATCTCTTAATTCAGATGGGAATGCTCAGGGTAGC	420
Db	1459	AACCTGAAGTACAGAGACAGATGAATCTCTTAATTCAGATGGGAATGCTCAGGGTAGC	1518
Qy	421	TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAAATCAGAAACT	480
Db	1519	TAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATGGATCTCCAGAAATCAGAAACT	1578
Qy	481	GAAAGAGTTCAATGACTGCTTACAAACACAGAGAAAGAAACAGGAAAATGGAGGAGA	540
Db	1579	GAAAGAGTTCAATGACTGCTTACAAACACAGAGAAAGAAACAGGAAAATGGAGGAGA	1638
Qy	541	GCCTCTTGGACCTGATCTTCAAGACCTTAAACCCCAAGTACAAACACATAAGGTGCTTCA	600
Db	1639	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTACAAACACATAAGGTGCTTCA	1698
2y	601	AGAAGATCTAGAACAGAACAAAGTCAAGGTCAATTTCTCTCCTCAGATGGTGGTAGT	660
Db	1699	AGAAGATCTAGAACAGAACAAAGTCAAGGTCAATTTCTCTCCTCAGATGGTGGTAGT	1758
2y	661	TGATGAATCTAGTGGAGATCACCCAACTGCTGCTTTGGAGAACACTTAAAGSTATTGGG	720
Db	1759	TGATGAATCTAGTGGAGATCACCCAACTGCTGCTTTGGAGAACACTTAAAGSTATTGGG	1818
2y	721	AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT	780
Db	1819	AGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACAT	1878
2y	781	CCCTCTCAATGGCAAGCTTCTACTGAAGACAGTGCCTTTTGTGCTATGGCTTTCAGA	840
Db	1879	CCCTCTCAATGGCAAGCTTCTACTGAAGACAGTGCCTTTTGTGCTATGGCTTTCAGA	1938
2y	841	AAAAGAGATGCAGTGAACAAAGATTCCACAACTGGCTTTAAAGATCAAAATGAATGTT	900
Db	1939	AAAAGAGATGCAGTGAACAAAGATTCCACAACTGGCTTTAAAGATCAAAATGAATGTT	1998
2y	901	ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	960
Db	1999	ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	2058
2y	961	GGCAAACTGTATTCTCAACAAAGATCTTCTTCAACACTGAAGATAAGTCAAGTGAC	1020
Db	2059	GGCAAACTGTATTCTCAACAAAGATCTTCTTCAACACTGAAGATAAGTCAAGTGAC	2118
2y	1021	CCAGAGACGGAAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAAA	1080
Db	2119	CCAGAGACGGAAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAAA	2178
2y	1081	ACTTGAAGAGGTACAGCACAGACT	1105
Db	2179	ACTTGAAGAGGTACAGCACAGACT	2203

Search completed: February 2, 2004, 17:38:19
 Job time : 716.863 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:11:40 ; Search time 129.879 Seconds
(without alignments)
7140.092 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000
Perfect score: 2101
Sequence: 1 98cgtatgctacacacagg.....acctcagcactcttggaagac 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Capext 1.0

Searched: 569378 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B-COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A-COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B-COMB.seq: *
5: /cgn2_6/prodata/2/ina/PTCUS-COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1131	53.8	5952	4	US-09-687-875A-1 Sequence 1, Appli
2	1103.4	52.5	13977	4	US-09-484-970B-60 Sequence 60, Appli
3	866.6	41.2	19307	3	US-08-836-022A-10 Sequence 10, Appli
4	866.6	41.2	19307	3	US-09-427-048A-10 Sequence 10, Appli
5	343	16.3	6045	4	US-09-091-501B-7 Sequence 7, Appli
6	343	16.3	10320	4	US-09-091-501B-9 Sequence 9, Appli
7	79.4	3.8	200	4	US-09-091-501B-5 Sequence 5, Appli
8	78.6	3.7	200	4	US-09-091-501B-4 Sequence 4, Appli
9	78.6	3.7	200	4	US-09-091-501B-6 Sequence 6, Appli
10	76.6	3.6	7218	1	US-08-232-463-14 Sequence 14, Appli
11	63.6	3.0	238	4	US-09-687-875A-13 Sequence 13, Appli
12	44	2.1	1230025	4	US-09-198-452A-1 Sequence 1, Appli
13	43.4	2.1	1179	4	US-09-107-532A-1186 Sequence 1186, Ap
14	40.4	1.9	2223	1	US-08-257-073-4 Sequence 4, Appli
15	39.2	1.9	16395	4	US-08-961-527-82 Sequence 82, Appli
16	38.6	1.8	1751	4	US-09-620-312D-847 Sequence 847, App
17	38.6	1.8	1995	1	US-08-425-069-3 Sequence 3, Appli
18	38.6	1.8	1995	2	US-08-317-844B-3 Sequence 3, Appli
19	38.2	1.8	1131	6	5180810-3 Patent No. 5180810
20	38.2	1.8	1784	6	5180810-2 Patent No. 5180810
21	38.2	1.8	1848	4	US-09-134-001C-447 Sequence 447, App
22	38.2	1.8	4529	4	US-09-620-312D-674 Sequence 674, App
23	38	1.8	1394	4	US-09-247-155-76 Sequence 76, Appli
24	37.4	1.8	2082	3	US-08-985-335-4 Sequence 4, Appli
25	37.4	1.8	2082	3	US-09-410-372-4 Sequence 4, Appli
26	37.2	1.8	2915	4	US-09-336-115C-5 Sequence 5, Appli
27	37.2	1.8	3902	4	US-08-961-527-212 Sequence 212, App

Patent No. 5210183
Sequence 22, Appli
Sequence 22, Appli
Sequence 17, Appli
Sequence 17, Appli
Sequence 14, Appli
Sequence 1, Appli
Sequence 236, App
Sequence 193, App
Sequence 13, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 24, Appli
Sequence 30, Appli
Sequence 5, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 38, Appli

28 36.8 1.8 1886 6 5210183-1
29 36.6 1.7 1845 4 US-08-887-534A-22
30 36.6 1.7 1845 4 US-09-527-431-22
31 36.2 1.7 289 3 US-09-007-005-17
32 36.2 1.7 289 3 US-08-244-796-17
33 36.2 1.7 2447 2 US-09-014-969-14
34 36.2 1.7 168575 4 US-09-426-290-1
35 36 1.7 608 3 US-08-630-915A-193
36 36 1.7 2873 4 US-09-461-474-13
37 35.8 1.7 790 3 US-08-139-937-12
38 35.6 1.7 4868 1 PCT-US93-11310-12
39 35.6 1.7 4858 5 PCT-US93-11310-12
40 35.6 1.7 7672 4 US-09-420-132-24
41 35.6 1.7 8257 4 US-09-595-684B-30
42 35.6 1.7 8789 1 US-08-328-254-5
43 35.6 1.7 10136 1 US-08-353-700-2
44 35.6 1.7 10136 5 PCT-US95-16216-2
45 35.2 1.7 3466 1 US-08-468-036-38

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICED PE
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/158,868
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match 53.8%; Score 1131; DB 4; Length 5952;
Best Local Similarity 58.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;
Qy 1 GAGCTATGCTACACACAGCTGCTTATGTCACCACTCTGACCCCTGACCCCTACACGAGCCATT 60
Db 891 GAGCTATGCTACACACAGCTGCTTATGTCACCACTCTGACCCCTGACCCCTACACGAGCCATT 950
Qy 61 TCCTTTCACAGCATTTGGAAGCTTCTGGAAGCAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
Db 951 TCCTTTCACAGCATTTGGAAGCTTCTGGAAGCAAGTCAATTTGGCAGTTCATTGATGGAGAG 1010
Qy 121 TGAAGTAAACCTGGAGCCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 180
Db 1011 TGAAGTAAACCTGGAGCCGTTATCAAAACAGCTTTAGAAGAAGTATTATCGTGGCTTCTTTTC 1070
Qy 181 TGCTGAGGACACATTTGGAAGCAAGGAGAGATTTCCTATGATGTGGAAGTGGTGAAGA 240
Db 1071 TGCTGAGGACACATTTGGAAGCAAGGAGAGATTTCCTATGATGTGGAAGTGGTGAAGA 1130
Qy 241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGCTTGG 300

Db 1131 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGGGTTGG 1190
Qy 301 TAAATTTCTACAAATGGGAGTAAAGCTGATTCGAAACAGCAAAATATATCAGAAGATGAAG 360
Db 1191 TAAATTTCTACAAATGGGAGTAAAGCTGATTCGAAACAGCAAAATATATCAGAAGATGAAG 1250
Qy 361 AACTGAAGTACAAAGACAGATGAATCTCTCTAAATTCAGATGGGAATCCCTCAGGGTAGC 420
Db 1251 AACTGAAGTACAAAGACAGATGAATCTCTCTAAATTCAGATGGGAATCCCTCAGGGTAGC 1310
Qy 421 TAGCATGGAATCAAGCAATTTACATAGATTTTAAATGATCTCCAGATCAGAACT 480
Db 1311 TAGCATGGAATCAAGCAATTTACATAGATTTTAAATGATCTCCAGATCAGAACT 1370
Qy 481 GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAACAGGAAATTCGAGGGAAGA 540
Db 1371 GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAACAGGAAATTCGAGGGAAGA 1430
Qy 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATTAAGGTGCTTCA 600
Db 1431 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACAAACATTAAGGTGCTTCA 1490
Qy 601 AGAAGATCTAGAAACAAAGACAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTGTAGT 660
Db 1491 AGAAGATCTAGAAACAAAGACAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTGTAGT 1550
Qy 661 TGATGAATCTAGTGAGATCAGCAACTGCTGTTTGAAGAAACAACTTAAGGTATTGGG 720
Db 1551 TGATGAATCTAGTGAGATCAGCAACTGCTGTTTGAAGAAACAACTTAAGGTATTGGG 1610
Qy 721 AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAGACAT 780
Db 1611 AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAGACAT 1670
Qy 781 CCTTCTCAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATGTCATGCTGCTTTCAGA 840
Db 1671 CCTTCTCAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTATGTCATGCTGCTTTCAGA 1730
Qy 841 AAAGAGATGAGTGAACAGATTTCAACAACTGGCTTTTAAAGATCAAATGAATGTT 900
Db 1731 AAAGAGATGAGTGAACAGATTTCAACAACTGGCTTTTAAAGATCAAATGAATGTT 1790
Qy 901 ATCAAGTCTTCAAAACTGGCGGTTTAAAGCGGATCTAGAAAGAAAGAAAGCAATCCAT 960
Db 1791 ATCAAGTCTTCAAAACTGGCGGTTTAAAGCGGATCTAGAAAGAAAGAAAGCAATCCAT 1850
Qy 961 GGGCAACTGTATTCACCTAAACAGATCTTCTTCAACACTGAAGAAATAGTCAGTGAC 1020
Db 1851 GGGCAACTGTATTCACCTAAACAGATCTTCTTCAACACTGAAGAAATAGTCAGTGAC 1910
Qy 1021 CCAGAAAGCGGAGCATGGCTGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1080
Db 1911 CCAGAAAGCGGAGCATGGCTGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAA 1970
Qy 1081 ACTTGAAGAGTACAGCACAG - 1102
Db 1971 ACTTGAAGAGTACAGCACAGCAAGAACTCAAATAGCAGTTCAAAGTAAACAAACCGGATGT 2030
Qy 1103 - 1102
Db 2031 GGAAGAGATTTGTCTAAAGGCGACATTTGTACAGGAAACACAGCCACTCAGCCAGT 2090
Qy 1103 - 1102
Db 2091 GAAAGGAAGTTAGAAGATCTGAGCTCTGAGTGGGAGGCGGTAAACCGTTTACTTCAAGA 2150
Qy 1103 - 1102
Db 2151 GCTGAGGGCAAGACGCTGACCTAGTCTCTGGACTGACCACTATATGAGGCTCTCTCTAC 2210
Qy 1103 - 1102
Db 2211 TCAGACTGTTACTCTGTGTGACACAACTGTGTGTTACTTAAGGAAACTGCCATCTCCAACT 2270

Qy 1103 - 1102
Db 2271 AGAAATGCCATCTTCTTGATGTTGGAGGTACCTGCTTGGCAGATTTTCAACCGGCTTG 2330
Qy 1103 - 1102
Db 2331 GACAGAACTTACCGACTGGCTTCTCTGCTTGATCAAGTTATAAAATCAGAGGGGTGAT 2390
Qy 1103 - 1102
Db 2391 GGTGGTGACCTTGGAGTATCAACGAGATGATCATCAAGCAGAGGCAACAAATGCAGGA 2450
Qy 1103 - 1102
Db 2451 TTTGGAACAGAGGCGTCCCGAGTTGGAAGAACTCATTTACCGCTGCCAAATTTGAAAAA 2510
Qy 1103 - 1102
Db 2511 CAAGACCAATCAAGAGGCTAGAACAAATCATTTACGGATCGAATTTGAAGAATTCAGAA 2570
Qy 1103 - 1102
Db 2571 TCAGTGGGATGAGTACAGAACACCTTCAGAACCGGAGCAACAGTTGAATGAATGTT 2630
Qy 1103 - 1102
Db 2631 AAAGGATTCACACAAATGGCTGGAAGCTAAGGAAGAGCTGAGCAGGTCTTTAGGACAGGC 2690
Qy 1103 - 1102
Db 2691 CAGGCCAAGCTTGAGTCATGGAAGAGGGTCCCTATACAGTAGATGCAATCCAAAGAA 2750
Qy 1103 - 1102
Db 2751 AATCAGAGAAACCAAGCAGTTGGCCCAAGACCTCCCGCAGTGGCAGACAAATGTAGATGT 2810
Qy 1103 - 1102
Db 2811 GGCAAATGACTTGGCCCTGAAACTTCTCGGGATTTATTTCTGCAGATGATACCAGAAAAAGT 2870
Qy 1103 - 1102
Db 2871 CCACATGATAACAGAGATATCAATGCCTCTTGGAGAAAGCATTCATAAAAGGGTGAGTGA 2930
Qy 1103 - 1140
Db 2931 GCGAGAGGCTGCTTTGGAAGAAACTCATAGATTACTGCAACAGATTCCCTCGGACCTGGA 2990
Qy 1141 AAAGTTTCTTGGCTGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1200
Db 2991 AAAGTTTCTTGGCTGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 3050
Qy 1201 CCGTAAGGAAGGCTCTTAGAAGACTCCAGGAGTAAAGAGCTGATGAACATGGA 1260
Db 3051 CCGTAAGGAAGGCTCTTAGAAGACTCCAGGAGTAAAGAGCTGATGAACATGGA 3110
Qy 1261 AGACCTCCAAAGTGAATTTGAAGCTCACAACAGATGTTTATCAACCTGGATGAAACAG 1320
Db 3111 AGACCTCCAAAGTGAATTTGAAGCTCACAACAGATGTTTATCAACCTGGATGAAACAG 3170
Qy 1321 CCAAAAATCTCTGAGATCCCTGGAGGTTCCGATGATGAGTCTCTGTTTACAAAGACGTTT 1380
Db 3171 CCAAAAATCTCTGAGATCCCTGGAGGTTCCGATGATGAGTCTCTGTTTACAAAGACGTTT 3230
Qy 1381 GGATACATGAACTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACATTTAGTCCCA 1440
Db 3231 GGATACATGAACTTCAAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACATTTAGTCCCA 3290
Qy 1441 TTTGGAAGCAGTCTCTGACCACTGGAAGCGTGTGACCTTTTCTCTGAGGAACTTTCTGGT 1500
Db 3291 TTTGGAAGCAGTCTCTGACCACTGGAAGCGTGTGACCTTTTCTCTGAGGAACTTTCTGGT 3350

1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTATTGGAGCGCACTTTCC 1560
3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTATTGGAGCGCACTTTCC 3410
1561 AGCAGTTACAGAGCAGAGATGATACATAGAGCCCTCAAGAGGGAATTTGAACCTTAAGA 1620
3411 AGCAGTTACAGAGCAGAGATGATACATAGAGCCCTCAAGAGGGAATTTGAACCTTAAGA 3470
1621 ACCTGTAAATCATGAGTACTCTTTGAGACTGTAGCAATATTTCTGACAGAGCAGCCCTTTGA 1680
3471 ACCTGTAAATCATGAGTACTCTTTGAGACTGTAGCAATATTTCTGACAGAGCAGCCCTTTGA 3530
1681 AGGACTAGAGAACTCTACAGAGGCCCGAGAGCTGCTCTCTGAGGAGAGAGCCCGAGAA 1740
3531 AGGACTAGAGAACTCTACAGAGGCCCGAGAGAGCTGCTCTCTGAGGAGAGAGCCCGAGAA 3590
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAATTTGAA 1800
3591 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAATTTGAA 3650
1801 CTTGCACTCCGCTGACTGCGCAGAGAAAATAGATGAGACCCCTTGAAGACTCTCCAGGAACT 1860
3651 CTTGCACTCCGCTGACTGCGCAGAGAAAATAGATGAGACCCCTTGAAGACTCTCCAGGAACT 3710
1861 TCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 1920
3711 TCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 3770
1921 CTTGCGAGCCGCTGGGCGATCTCCTCAITGACTCTCTCAAGATCACTTCGAGAAAGTCAA 1980
3771 CTTGCGAGCCGCTGGGCGATCTCCTCAITGACTCTCTCAAGATCACTTCGAGAAAGTCAA 3830
1981 GCACCTTCGAGGAGAAATTCGCGCTCTGAAGAGAACTGAGCCACGTCATGACCTTGC 2040
3831 GCACCTTCGAGGAGAAATTCGCGCTCTGAAGAGAACTGAGCCACGTCATGACCTTGC 3890
2041 TCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTTCAGCACTCTGGAAGA 2100
3891 TCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTTCAGCACTCTGGAAGA 3950
2101 C 2101
3951 C 3951

RESULT 2

US-09-484-970B-60

Sequence 60, Application US/09484970B

Patent No. 6426186

GENERAL INFORMATION:

APPLICANT: Jones, Karen A.

APPLICANT: Volkmut, Wayne

APPLICANT: Walker, Michael G.

TITLE OF INVENTION: BONE REMODELING GENES

FILE REFERENCE: PB-0014 US

CURRENT APPLICATION NUMBER: US/09/484,970B

CURRENT FILING DATE: 2000-01-18

NUMBER OF SEQ ID NOS: 172

SOFTWARE: PERL Program

SEQ ID NO 60

LENGTH: 13977

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1

NAME/KEY: unsure

LOCATION: 11721-11761, 12294, 13969

OTHER INFORMATION: a, t, c, g, or other

US-09-484-970B-60

Query Match

Best Local Similarity

52.5%; Score 1103.4; DB 4; Length 13977;

99.9%; Pred. No. 0;

Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCTACACACAGCGTCTTATGTACCACCTCTGACCCCTACACGAGGCCCAATT 60
DB 1099 GAGCTATGCTACACACAGCGTCTTATGTACCACCTCTGACCCCTACACGAGGCCCAATT 1158
QY 61 TCCTTCACAGCAATTTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTCATTTGATGAGAG 120
DB 1159 TCCTTCACAGCAATTTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTCATTTGATGAGAG 1218
QY 121 TGAAGTAAACCTCGGACCGTTATCAACACGCTTTAGAAAGATTAATTCGTGGCTTTCTTC 180
DB 1219 TGAAGTAAACCTCGGACCGTTATCAACACGCTTTAGAAAGATTAATTCGTGGCTTTCTTC 1278
QY 181 TGTGAGGACACATTTCAAGCAAGGAGAGATTTCTAATGATGTGAAAGTGGTGAAGA 240
DB 1279 TGTGAGGACACATTTCAAGCAAGGAGAGATTTCTAATGATGTGAAAGTGGTGAAGA 1338
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGGCCCATCAGGCGCGGTGG 300
DB 1339 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGGCCCATCAGGCGCGGTGG 1398
QY 301 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAATCAAGA 360
DB 1399 TAATATTCTACAAATTTGGGAAGTAAGCTGATTGGAAACAGGAAAATTTATCAGAAATCAAGA 1458
QY 361 AACTGAGTACAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCTCAGGGTAGC 420
DB 1459 AACTGAGTACAGAGCAGATGAATCTCCTAAATCAAGATGGGAATGCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAATCT 480
DB 1519 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAATCT 1578
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGAAAATCGAGGAAGA 540
DB 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAAGAAAATCGAGGAAGA 1638
QY 541 GCCTCTTGACCTGATCTTTGAAGACTTAAGCCCAAGTACAAACACATTAAGGTCCTTCA 600
DB 1639 GCCTCTTGACCTGATCTTTGAAGACTTAAGCCCAAGTACAAACACATTAAGGTCCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGAAACAAGTCAGGGTCAATTTCTCACTCACTGCTGGTGGTAGT 660
DB 1699 AGAAGATCTAGAACAAAGAAACAAGTCAGGGTCAATTTCTCACTCACTGCTGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTCTTTGGAGAACAACTTAAGGTATTGGG 720
DB 1759 TGATGAATCTAGTGGAGATCAGCAACTGCTCTTTGGAGAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 780
DB 1819 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 CCTTCTCAAAATGGCAACGCTTTACTAGAAACAGTGCCTTTTTAGTGCATGCTTTTCA 840
DB 1879 CCTTCTCAAAATGGCAACGCTTTACTAGAAACAGTGCCTTTTTAGTGCATGCTTTTCA 1938
QY 841 AAAAGAGATGCAAGTGAACAAAGATTCACACACTGGCTTTAAAGATCAAAATGAATGTT 900
DB 1939 AAAAGAGATGCAAGTGAACAAAGATTCACACACTGGCTTTAAAGATCAAAATGAATGTT 1998
QY 901 ATCAAGTCTTCAAAAATCGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
DB 1999 ATCAAGTCTTCAAAAATCGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGATTTCACTCAAAACAGATCTTTTCAACACTGAGAGTAAGTCACTGAGTGC 1020
DB 2059 GGGCAAACTGATTTCACTCAAAACAGATCTTTTCAACACTGAGAGTAAGTCACTGAGTGC 2118
QY 1021 CCAGAAAGACGAGATGCTGGATAAATTTCCCGCGTGTGGGATAATTTAGTCCAAAA 1080
DB 2119 CCAGAAAGACGAGATGCTGGATAAATTTCCCGCGTGTGGGATAATTTAGTCCAAAA 2178

QY 1081 ACTTGAAAGAGTACAGCAGACT 1105
DB 2179 ACTTGAAAGAGTACAGCAGATT 2203

RESULT 3

US-08-836-022A-10/c
; Sequence 10, Application US/08836022A
; Patent No. 6001557
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,022A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,381
; FILING DATE: 28-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVFN.008PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
US-08-836-022A-10

Query Match 41.2%; Score 866.6; DB 3; Length 19307;
Best Local Similarity 86.5%; Pred. No. 6.1e-253;
Matches 956; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 1 GAGCTATGCTACACAGGCTGTATGTCCACACCTCTGACCCCTACACGAGCCCAT 60
DB 13568 GAGTTATGCTTCCACAGGCTGTATGTGCGCACTCTGATTCACAGAGCCCTA 13509
QY 61 TCCTTCACAGATTGGAGCTCCTGAAGACAGATCATTTGGCAGTTCATTGATGGAGAG 120
DB 13508 TCCTTCACAGATTGGAGCTCCTGAAGACAGATCATTTGACAGTTCATTGATGGAGAG 13449
QY 121 TGAAGTAACTGGACCGCTTATCAACAGCTTTAGAAAGATATATCGTGGCTTCTTC 180
DB 13448 GGAAGTAACTCTGGATGTACCAACCTGCTTTAGAGAGATCTTCATGGCTTCTTC 13389
QY 181 TGCTGAGGACATTCGAAAGCAGAGGAGATTTCTTAATGATGTGAAGTGTGAAGA 240
DB 13388 TGCCGAGGATACATTCGAGAGCAGAGAGATTTCAAAATGATGTTGAAGAGTGAAGA 13329

QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCCGGTTGG 300
DB 13328 ACAGTTTCATCTCATGAGGGATTCATGATGGATTTGACATCTCATCAGGACTTGTGG 13269
QY 301 TAAATTTCTCAATTTGGGAAGTAAAGCTGATTGGAAACAGGAAAAATTTATCAGAAATGAAGA 360
DB 13268 TAAATTTCTCAAGTTAGGAAGTCAACTAGTTGGAAAAAGGAAAAATTTATCAGAAATGAAGA 13209
QY 361 AACTGAAGTACAGAGCAGATGATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
DB 13208 AGCTGAAGTCAAGAGCAAAATGATCTCTAAATTCAGATGGGAATGTCTCAGGGTAGC 13149
QY 421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGATCAGAACT 480
DB 13148 TAGCATGAAAAACAAAGCAATTTACACAAAGTTCTAAATGGATCTCCAGATCAGAACT 13089
QY 481 GAAAGAGTTGAATGACTGCTCTAACAAAAACAGAAAGAACCAAGGAAAAATGGAAGGAAGA 540
DB 13088 AAAAGAACTAGATGACTGTGTTTAAACAAAACCTGAAGAGAGAACTAAGAAATGGAAGGAAGA 13029
QY 541 GCCTCTTCGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAAACATAAGGTGCTTCA 600
DB 13028 GCCCTTTGGACCTGATCTTTGAAGATCTTAAATGCAAGTACAAACATAAGGTGCTTCA 12969
QY 601 AGAAGATCTAGAACAGAACAGTCAAGTCAAGTCAATCTCTCACTCAGATGGTGTGCTAGT 660
DB 12968 AGAAGATCTAGAACAGAGGAGGTCAGGTCAGGTCACCTCGCTCACTCAGATGGTGTGCTAGT 12909
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGAAGAAACAACTTAAAGTATTGGG 720
DB 12908 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTTTGAAGAAACAACTTAAAGTATTGGG 12849
QY 721 AGATCGATGGGCAAAACATCTCTAGATGGAAGAGAGAGCGCTGGTCTCTTTTACAGACAT 780
DB 12848 AGATCGATGGGCAAAATATCTCGATGAGTCAAGACCGCTGGATTTGTTTACAGATAT 12789
QY 781 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTAAAGTGCATGCTTTTACAGA 840
DB 12788 TCTTCTAAATGGCAGCATTTTACTGAGAACAGTGCCTTTTAAAGTGCATGCTTTTACAGA 12729
QY 841 AAAAGAGATGAGTGAACAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
DB 12728 AAAAGAGATGCAATGAAGAACATTCAGACAAGTGGCTTTAAAGATCAAAATGAAATGAT 12669
QY 901 ATCAAGTCTTCAAAAACTGGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 960
DB 12668 GTCAAGTCTTCAAAAAATATCTCTTTTAAATAGATCTAGAAAAGAAAAAGCAATCCAT 12609
QY 961 GGGCAAACTGTATTCACTCAAAACAGATCTTTCTTTCAACACTGAAGATTAAGTCAGTGAC 1020
DB 12608 GGAATAAACTAAGTTCACTCAATCAAGATCTACTTTGCGCACTGAAAAATTAAGTCAGTGAC 12549
QY 1021 CCAGAGAGCGGAGCATGGCTGGATTAAGTTCGCGGTTGGGATTAATTTAGTCCAAAA 1080
DB 12548 TCAAAAGATGGAATCTGGATGGAACCTTTGCAACAGTTGGGACAAATTTTAAACCCAAA 12489
QY 1081 ACTTGAAAGAGTACAGCAGACT 1105
DB 12488 ACTTGAAAGAGTTTCAGCACAATA 12464

RESULT 4

US-09-427-048A-10/c
; Sequence 10, Application US/09427048A
; Patent No. 6203975
; GENERAL INFORMATION:
; APPLICANT: Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Fisher, Krishna J.
; APPLICANT: Chen, Shu-Jen
; APPLICANT: Weitzman, Matthew
; TITLE OF INVENTION: Improved Adenovirus Virus and
; Methods of Use Thereof

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999

PRIOR APPLICATION DATA:
CLASSIFICATION: <Unknown>
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVEN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 10:
IS-09-427-048A-10

Query Match 41.2%; Score 866.6; DB 3; Length 19307;
Best Local Similarity 86.5%; Pzed. No. 6.1e-253;
Matches 956; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
2Y 1 GAGCTATGCTCACACAGGCTGCTTATGTCACCACTCTGACCCCTACCGAGGCCCAATT 60
2B 13568 GAGTTATGCTTCCACACAGGCTGCTTATGTCACCACTCTGACCCCTACCGAGGCCCAATT 13509
2Y 61 TCCTTCACAGCATTTGGAAGCTCCCTGACACAGTCATTTGGCAGTTCATTGATGGAGAG 120
2B 13508 TCCTTCACAGCATTTGGAAGCTCCCTGACACAGTCATTTGGCAGTTCATTGATGGAGAG 13449
2Y 121 TGAAGTAAACCTGGACCGCTTATCAACACAGCTTTAGAAGAGTATTATCGTGGCTTCCTTC 180
2B 13448 GGAAGTAAATCTGGATAGTTACCAAACTGCTTTAGAAGAGTATTATCGTGGCTTCCTTC 13389
2Y 181 TGCTGAGGACATTCGACGACACAGGAGAGATTTCTATGATGGAAGTGGTGAAGA 240
2B 13388 TGCCGAGGATACATTCGAGGACACAGGAGAGATTTCAAATGATGATGGAAGTGAAGA 13329
2Y 241 CCAGTTTTCATCTCATGAGGCTCATGATGATTTGACGCCCATCAGGCCCGGTTGG 300
2B 13328 ACAGTTTTCATCTCATGAGGCTCATGATGATTTGATGATGATCTCATCAGGACTTGTGG 13269
2Y 301 TAATATCTCAATTTGGGAAGTAACTGATGATGGAACAGGAAATTTATCAGAAGATGAAGA 360
2B 13268 TAATGTTCTACAGTTAGGAAGTCAACTAGTTGGAAGAGGAAATTTATCAGAAGATGAAGA 13209
2Y 361 AACTGAAGTACAGGACGAGATCACTCTCAATTTCAAGATGGGAATCCCTCAGGCTAGC 420
2B 13208 AGCTGAAGTGAAGAACAAATGAAATCTCTCAATTTCAAGATGGGAATCTCTCAGGCTAGC 13149
2Y 421 TAGCATGGAACCAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAACT 480
2B 13148 TAGCATGGAACCAAGCAATTTACACAGAGTTCTTAATGGATCTCCAGAAATCAGAAATT 13089

QY 481 CAAAGACTTGAATGACTGCTAAACAAAACAGAAAGAAAGAAACAAAGGAAATGAGGAAGA 540
DB 13088 AAAAGAACTAGATGACTGGTTTAAACAAAACCTGAAGAGAACTAAAGAAATGAGGAAGA 13029
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACCCGCAAGTACAAACAATAGGTGCTTCA 600
DB 13028 GCCTCTTGGACCTGATCTTGAAGATCTTAAATGCCAAGTACAAACAATAGGTGCTTCA 12969
QY 601 AGAAGATCTAGAAACAAGAAACAAGTCAAGGCTCAATTTCTCCTCCTCAGATGCTGCTGCTAGT 660
DB 12968 AGAAGATCTAGAAACAAGACAGGCTCAGGCTCACTCGCTCCTCCTCAGATGCTGCTAGT 12909
QY 661 TGATGAATCTAGTGGAGATCACCGCACTGCTGCTTGGAAAGAAACAATTAAGTATTTGG 720
DB 12908 TGATGAATCCAGCGGTGATCATGCAACAGCTGCTTGGAAAGAAACAATTAAGTATTTGG 12849
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAAAGCCGCTGGGTTCTTTTACAAGACAT 780
DB 12848 AGATCGATGGGCAATATCTGAGATGGACTGAAGACCGCTGGATTTGTTTACAAGATAT 12789
QY 781 CCTTCTCAATGGCAACGCTTCTTACTGAAGAACAGTCCCTTTTGTGCTGCTGCTTTTCA 840
DB 12788 TCCTTCAATGGCAACGCTTCTTACTGAAGAACAGTCCCTTTTGTGCTGCTGCTTTTCA 12729
QY 841 AAAAGAGATGCAAGTGAACAAGATTCACACCACTGGCTTTAAGATCAAAATGAAATGTT 900
DB 12728 AAAAGAGATGCAATGAAGAACATTCAGCAAGTGGCTTTTAAAGATCAAAATGAAATGAT 12669
QY 901 ATCAAGTCTTCAAAAACCTGGCCCTTTTAAAGCGGATCTAGAAAAGAAAGCAATTCCTAT 960
DB 12668 GTCAAGTCTTCAAAAATATCTACTTTAAATAATAGATCTAGAAAAGAAAGCAATTCCTAT 12609
QY 961 GGGCAACTGTTTCTACTCAACCAAGATCTTTTCAACACTGGAAGAAATTAAGTCACTGAC 1020
DB 12608 GSAAGAACTAAGTTCACTCAATCAAGATCTACTTTTGGCACTGAAATTAAGTCACTGAC 12549
QY 1021 CCAGAAGACGGAAGCATGGCTGGATTAATTTGGCCGGTGTGGGATAATTTAGTCCAAA 1080
DB 12548 TCAAGAGATGGAATCTGATGGAATCTTGAACAACGTTGGACAACTTTAACCCAAA 12489
QY 1081 ACTGAAAAGAGTACAGCACAGACT 1105
DB 12488 ACTGAAAAGAGTTCAGCAAAATT 12464

RESULT 5
US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(6037)

Query Match	16.3%	Score 343;	DB 4;	Length 10320;
Best Local Similarity	59.0%;	Prod. No. 1.5e-93;		
Matches 589;	Conservative 0;	Mismatches 410;	Indels 0;	Gaps 0;
QY	104	AGTTCATTGATGAGAGTGAAGCTAAACCTGGACCGCTTATCAACAGAGCTTTAGAGAGAAGTA	163	
DB	914	AGCACCGCTCACTGAAGTGGACATGGATTTGGACAGCTACCCAGATAGCGCTAGAGAAATGG	973	
QY	164	TTATCTCGGGCTTTCTTCTGCTGAGGACACATTCGAAGCACAAAGGAGAGATTTCTTAATGAT	223	
DB	974	CTGACGCTGGCTGCTGTCGCCGGGAGACAGTTCACGAGCAGATGACATTTCTTGATGAT	1033	
QY	224	GTGAAAGTGTGAAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCC	283	
DB	1034	GTGAAAGAAGTCAAAAGACGAGTTGTGTAACCATGAAACTTTTATGATGAGAGCTGACAGCA	1093	
QY	284	CATCAGGGCCGGCTTGGTAAATTCTACAAATTTGGGAAGTAAAGCTGATTTGGAAACAGGAAAA	343	
DB	1094	CACCAGAGCAGCGTGGGGAGCGTCTCGAGGCTGGCAACCCAGCTGATGACACAAGGGACT	1153	
QY	344	TTATTCAGAAGATGAAGAAACTGAAGTACAAAGAGCAGATGAATCTCTAAATTTCAAGATGG	403	
DB	1154	CTGTCAGAGGAGGAGTTTGATATCCAGGAAACAGATGACCTTGCTGAATGCAAGGTGG	1213	
QY	404	GAATGCCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGAT	463	
DB	1214	GAGCGGCTCCGGGTGGAGAGCATGAGAGAGGCACTCCCGGCTGCACGAGCGTCTGATGGAG	1273	

QY 464 CTCAGAAATCAGAAACTGAAAGAGTGAATGACTGCTTAACAAACACAGAAAGAAACA 523
Db 1274 CTCAGAAAGAAACAGCTGCAGCAGCTCTCAAGTGTGCTGCGCTCACAGAGAGCGCATT 1333
QY 524 AGGAAATGAGGAAAGAGCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCCAAGTACAA 583
Db 1334 CAGAAGTGGAGAGCCTCGCTGGGTGATGACCTGCCCTCCCTGCAAGTGTCTTCAA 1393
QY 584 CAACATAAGTGTCTCAAGAAATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTCACT 643
Db 1394 GAACATAAAAGTTTGCAAAATGACCTTGAAGCTGAACAGGTGAAGTAAATTCCTTAAC 1453
QY 644 CACATGTGTGTGATGTGATGATCTAGTGGAGATCAGCACTGCTCTTTGGAAGAA 703
Db 1454 CACATGTGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1513
QY 704 CAACCTAAGGTATTTGGAGATCGATGGCAACATCTGTAGATGACAGAGACCGCTGG 763
Db 1514 CAGTTACAGAACTGGGTGAGCGCTGGACAGCTGTATGCCGTGGACTGAAGACGTTGG 1573
QY 764 GTTCTTTACAGACATCTTCTCAATGGCAACGCTTACTGAGAAACAGTGCCTTTT 823
Db 1574 AACAGTGTCAAGAAATCAGTATCTGTGGCAGGAATATTGGAAGAGCAGTGTCTGTG 1633
QY 824 AGTGCATGGCTTTTCAAGAAAGAGATGAGTGAACAGATTCACACAACTGGCTTTAA 883
Db 1634 GAGGCTTGGCTCACCGAAGAGAGAGGCTTTGATTAAGTTCAACACGAACTTTAA 1693
QY 884 GATCAAAATGAATGTATCAAGTCTTCAAACTGGCGCTTTTAAAGCGGATCTAGAA 943
Db 1694 GACCAAGAGAACTAAGTGTCAAGTGTCCGCGCTGTGCTATATTGAAGAGACATGGA 1753
QY 944 AAAAAAAGCAATCCATGGGCAACTGTATCACTCAACAAAGATCTTCTTCAACACTG 1003
Db 1754 ATGAAGAGGAGACTGTGATCACTGAGTGAATGGCCAGATGTGGGCCAATATCTC 1813
QY 1004 AGAATAAGTCACTGACCCAGAGACGGAAGCATGCTGATTAACCTTTGCCCGGTGTTGG 1063
Db 1814 AGTAATCCCAAGCACTTAAGAGATGAACAGTGACTCTGAGGAGCTAACACAGAGATGG 1873
QY 1064 GATAATTTAGTCCAAAACCTTGAAGAGATACAGACAG 1102
Db 1874 GATCTCTGTTTCAAGAGACTCGAAGACTCTTCTTAACCAAG 1912

RESULT 7

US-09-091-501B-5
; Sequence 5, Application US/09091501B
; Patent No. 6518413

; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 200
; TYPE: DNA

US-09-091-501B-5

Query Match

3.8%; Score 79.4; DB 4; Length 200;

Best Local Similarity 64.3%; Pred. No. 2.4e-14;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 565 CCTAAAGCGCCAACTCAACACATTAAGGTGCTTCAGAAAGATCTAGAAACAAAGCAACT 624
Db 16 CCTCAAAACCTGCTTGAAGAACATAAAAGTTTGCAAAAGTGACCTCGAAGCTGAGCAGGT 75
QY 625 CAGGTCAATTTCTCTCACTCACATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 684
Db 76 GAAGGTGAATTCCTTAACCTCATATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 135
QY 685 AACTGCTGCTTTTGAAGAAACAACTTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAG 744
Db 136 CACAGCTGTTTGAAGATCAGTTACAGAAACTGGGTGAGCGCTGGACAGCTGTATGCCG 195
QY 745 ATGGA 749
Db 196 CTGGA 200

RESULT 8

US-09-091-501B-4
; Sequence 4, Application US/09091501B
; Patent No. 6518413

; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 200
; TYPE: DNA

US-09-091-501B-4

Query Match 3.7%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 4.2e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 553 TGATCTTGAAGACCTAAAGCGCCAACTCAACACATTAAGGTGCTTCAGAAAGATCTAGA 612
Db 4 TGACCTGCTCCCTCGCAGAGCTGCTTCAAGAACATAAAAGTTTCAAAATGACCTTGA 63
QY 613 ACAAGAACAACTCAGGCTCAATTTCTCTCACTCACTGCTGGTGGTGGTGGTGGTGGT 672
Db 64 AGCTGAACAGTCAAGGTAAATTCCTTAACCTCACTGCTGGTGGTGGTGGTGGTGGTGG 123
QY 673 TGGAGATCAGCAACTGCTGCTTTGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 732
Db 124 TGGGAGAGTGCACAGCTCTTCTGGAAGATCAGTTTACAGAACTGGGTGAGCGCTGGAC 183
QY 733 AAACATCTGTAGATGGA 749
Db 184 AGCTGTATGCCGTGGA 200

RESULT 9

US-09-091-501B-6
; Sequence 6, Application US/09091501B
; Patent No. 6518413

; GENERAL INFORMATION:

US-09-687-875A-13

Query Match 3.0%; Score 63.6; DB 4; Length 238;
Best Local Similarity 94.3%; Pred. No. 1.7e-09;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1348 TTCGATGATGACGCTCTGTTACAAAGAGCTTTGGATACATGAACCTTCAAGTGAGTGA 1407
DB 169 TTGGAGCGACGCGACTGTTTACAAAGAGCTTTGGATACATGAACCTTCAAGTGAGTGA 228

QY 1408 ACTTCGAAA 1417
DB 229 ACTTCGAAA 238

RESULT 12

JS-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (15001)-(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (30001)-(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (45001)-(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (60001)-(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (75001)-(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (90001)-(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (105001)-(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (120001)-(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (135001)-(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (150001)-(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (165001)-(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (180001)-(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (195001)-(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (210001)-(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (225001)-(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (240001)-(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (255001)-(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (270001)-(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (285001)-(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (300001)-(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (315001)-(330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (330001)-(345000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (345001)-(360000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (360001)-(375000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (375001)-(390000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (390001)-(405000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (405001)-(420000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (420001)-(435000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (435001)-(450000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (450001)-(465000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (465001)-(480000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (480001)-(495000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (495001)-(510000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (510001)-(525000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (525001)-(540000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (540001)-(555000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (555001)-(570000)

OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature

Query Match

2.1%; Score 44; DB 4; Length 1230025;

Best Local Similarity 53.5%; Pred. No. 0.5;
Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 342 AATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAATTCAGAT 401
Db 656477 AATTAGAAGAGAGAGAGAGAGAAATAGGATATCAAGACTCAGATCAAAAT 656418
QY 402 GGAATGCTCAGGCTAGCTAGCATGGAAAAACAAGCAATTTACATAGATTTTAATGG 461
Db 656417 GGGTTTCGATCACTCAAGCTGCTAAATTACATAACGTCAGTACAGCAATTTATGTGG 656358
QY 462 ATCTCAGATCAGAACTGAAGAGTTGAATGCTAGCTGGCTAACAAAACAGA 513
Db 656357 CAATTAAGCAGAAAAAACTAAAGAGTTCTTAAGAGACGCGCTGGGAATAGA 656306
RESULT 13
US-09-107-532A-1186
; Sequence 1186, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1179 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...1179
; SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186

Query Match
Best Local Similarity 50.7%; Pred. No. 0.0073;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

2Y 471 ATCAGAACTGAAAGAGTTGAATGCTGCTAACAAAAACAGAGAAACAAAGAAAA 530
Db 602 AGCAAAATAAAGATTTGCTGATCAGACAAAGAAAAATGGAGATACGATCGGAGAA 661
2Y 531 TGGAGGAAGAGCTCTTGGACCTGATCTTGAAGACTTAAACCGCAAGTACACACATA 590
Db 662 TTGTAGAGTGTCTGTGGAGGGTTCCAGCTGGATTGGAAGCTTACGTACATGGGACA 721
2Y 591 AGTGCTTCAAGAAGATCTAGAACAGAAACAAGTCAGGCTCAATCTCTCACTCACATGG 650
Db 722 CGAAGCTAGATGCCAAATCGCACAAAGCTGTGGTAGTATCAATGCCTTTAAAGCGTAG 781
2Y 651 TGGTGCTAGTGTGATGATCTAGTGG 675
Db 782 AATTGGGTGCGATTCACTTCTGG 806

RESULT 14
JS-08-257-073-4
; Sequence 4, Application US/08257073
; Patent No. 5766597
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: de Taisne, Charles
; APPLICANT: Tine, John A.
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 03-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2223 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

JS-08-257-073-4
Query Match 1.9%; Score 40.4; DB 1; Length 2223;
Best Local Similarity 51.7%; Pred. No. 0.092;
Matches 92; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
Y 468 AGAATCAGAACTGAAAGAGTTGAATGCTGCTAACAAAAACAGAGAAACAAAGAA 527

Db 2033 AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 2092
QY 528 AAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAGTACACAC 587
Db 2093 AAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 2152
QY 588 ATAAGGTCTTCAAGAGATCTAGAACAAAGAACAGTCAGGGTCAATTTCTCTCACTCA 645
Db 2153 AAGAAGAAAGAAACAAGAAAGAAAGAAAGAAAGAAATAGTACCAGAAAATTTGACAACTGA 2210

RESULT 15
US-08-961-527-82
; Sequence 82, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16995 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-82

Query Match 1.9%; Score 39.2; DB 4; Length 16995;
Best Local Similarity 50.0%; Pred. No. 0.83;
Matches 98; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 417 TAGCTAGCATGGAACAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGA 476
Db 4967 TAGTGAATTTCTATTAAATATGGTTTCCATATTGATTTTGTGGATGCAATATCAGA 5026
QY 477 AACTGAAAGAGTTGAATGATCGCTTAACAAAAACAGAAAGAAAGAAACAAAGAAAAATGGAGG 536
Db 5027 TAATTCATGCTCTGAATGAAAGGGCAAGTATGAAATTTGAAATTCACAGAAAGGACAGT 5086
QY 537 AAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCCAAGTACACACATAGGTGC 596
Db 5087 CTCGTATAGAGTAACATAATGTGGTGGCATTAAAGCAAGACATCTGTGAGCAAGTCA 5146
QY 597 TTCAAGAAAGATCTAGA 612
Db 5147 TTGAAGCGATTCAAAA 5162

Search completed: February 2, 2004, 06:15:45
Job time : 145.879 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 499.045 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000

Perfect score: 2101

Sequence: 1 gagctatgctacacagg.....acctcagcactctggaagac 2101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq.19Jun03.*
1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2101	100.0	4182	24	Human dystrophin m
2	2101	100.0	5149	24	Adeno-associated v
3	1579.8	75.2	3999	24	Human dystrophin m
4	1579.8	75.2	4966	24	Adeno-associated v
5	1579.8	75.2	4990	24	Human dystrophin m
6	1443	68.7	3858	24	Adeno-associated v
7	1443	68.7	4825	24	Adeno-associated v
8	1443	68.7	4848	24	Adeno-associated v

9	1443	68.7	5060	24	Adeno-associated v
10	1293	61.5	8689	24	DNA encoding mini-
11	1131	53.8	5952	22	Human dystrophin g
12	1114	53.0	4414	24	Adeno-associated v
13	1103.4	52.5	5417	24	DNA encoding mini-
14	1103.4	52.5	11241	24	cDNA encoding huma
15	1103.4	52.5	11443	24	DNA encoding mini-
16	1103.4	52.5	12923	10	Sequence of human
17	1103.4	52.5	13957	24	Human breast cancer
18	1103.4	52.5	13957	24	Human dystrophin g
19	1103.4	52.5	13957	24	Gene #2284 used to
20	1103.4	52.5	13957	24	cDNA encoding huma
21	1103.4	52.5	13957	24	Human bone remodel
22	1102	52.5	3446	24	Human dystrophin m
23	1100.2	52.4	11058	24	Human dystrophin p
24	1097.8	52.3	1991	24	Human dystrophin N
25	999	47.5	2169	24	Human dystrophin r
26	999	47.5	3531	24	Human dystrophin m
27	999	47.5	4498	24	Adeno-associated v
28	997	47.5	5339	24	DNA encoding mini-
29	996	47.4	5462	24	DNA encoding mini-
30	911.6	43.4	4402	21	A rod shortened dy
31	866.6	41.2	13815	24	cDNA encoding mous
32	866.6	41.2	13815	24	Mouse ischaemic co
33	866.6	41.2	13307	17	Shuttle vector pAd
34	858.6	40.9	13815	19	Mus musculus dystr
35	855.4	40.7	3275	10	Partial sequence o
36	787	37.5	3510	24	Human dystrophin m
37	787	37.5	4476	24	Adeno-associated v
38	777	37.0	1667	24	Human dystrophin N
39	718.8	34.2	4075	21	A rod shortened dy
40	714.6	34.0	4402	21	A rod shortened dy
41	652	31.0	1821	24	Human dystrophin r
42	450	21.4	1340	24	Human dystrophin N
43	409.2	19.5	3747	21	A rod shortened dy
44	387	18.4	387	24	Human dystrophin s
45	361.4	17.2	10302	24	Human utrophin gen

ALIGNMENTS

RESULT 1

AAD37230
ID AAD37230 standard; DNA; 4182 BP.

XX AAD37230;

XX 21-AUG-2002 (first entry)

XX Human dystrophin minigene delta4173.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.

XX Homo sapiens.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular

PT. dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

XX Example 1; Page 43-44; 71pp; English.

XX CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta4173 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8059-10227 (rods R22, R23 and R24,
CC hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX SQ Sequence 4182 BP; 1309 A; 927 C; 970 G; 976 T; 0 other;

Query Match 100.0%; Score 2101; DB 24; Length 4182;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAGCTATGCTTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT	60
DB	900	GAGCTATGCTTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACACGGAGCCCAT	959
QY	61	TCCTTTCACAGCATTTGGGAAGCTCTGAAGACAGTCAATTTGGCAGTTCATTTGATGGAGAG	120
DB	960	TCCTTTCACAGCATTTGGGAAGCTCTGAAGACAGTCAATTTGGCAGTTCATTTGATGGAGAG	1019
QY	121	TGAAGTAAACCTGGACCTTATCAACAGCTTTAGAGAGTATTTATCGTGGCTTCTTTC	180
DB	1020	TGAAGTAAACCTGGACCTTATCAACAGCTTTAGAGAGTATTTATCGTGGCTTCTTTC	1079
QY	181	TGCTGAGGACACATTTGGAAGCTGGAAGACAGGAGAGATTTCTTAATCATGTGGAAGTGAAGA	240
DB	1080	TGCTGAGGACACATTTGGAAGCTGGAAGACAGGAGAGATTTCTTAATCATGTGGAAGTGAAGA	1139
QY	241	CCAGTTTCACTATGAGGGGTATCATGATGATTTGACAGCCCATCAGGGCCGGGTGG	300
DB	1140	CCAGTTTCACTATGAGGGGTATCATGATGATTTGACAGCCCATCAGGGCCGGGTGG	1199
QY	301	TAATATTTCTACAAATTTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	360
DB	1200	TAATATTTCTACAAATTTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	1259
QY	361	AACTGAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGATGCTTCAGGGTAGC	420
DB	1260	AACTGAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGATGCTTCAGGGTAGC	1319
QY	421	TAGCTGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATTCAGAACT	480
DB	1320	TAGCTGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATTCAGAACT	1379
QY	481	GAAAGAGTTGAATGCTGCTTAAACAAAGCAAGAAAGCAAGAAAGTGAAGAGAGAGAGAGAGAG	540
DB	1380	GAAAGAGTTGAATGCTGCTTAAACAAAGCAAGAAAGCAAGAAAGTGAAGAGAGAGAGAGAGAG	1439
QY	541	GCTCTTGACCTGATCTTGAAGACTTAAACGCGCAAGTACAAACAACTAAAGGTGCTTCA	600
DB	1440	GCTCTTGACCTGATCTTGAAGACTTAAACGCGCAAGTACAAACAACTAAAGGTGCTTCA	1499
QY	601	AGAGATCTAGAACAAAGACAGTCAAGGTCAATTTCTCTCATCTCATGCTGGTGGTGGT	660
DB	1500	AGAGATCTAGAACAAAGACAGTCAAGGTCAATTTCTCTCATCTCATGCTGGTGGTGGT	1559
QY	661	TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAACCACTTAAGGTATTTGG	720
DB	1560	TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAACCACTTAAGGTATTTGG	1619

QY	721	AGATCGATGGGCAACACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAGACAT	780
DB	1620	AGATCGATGGGCAACACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAGACAT	1679
QY	781	CTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTAAAGTCAAGTGGCTTTTCA	840
DB	1680	CTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTAAAGTCAAGTGGCTTTTCA	1739
QY	841	AAAGAGATCGAGTGAACAGAGATTCACAACTGGCTTTTAAAGTCAAAATGAATGTT	900
DB	1740	AAAGAGATCGAGTGAACAGAGATTCACAACTGGCTTTTAAAGTCAAAATGAATGTT	1799
QY	901	ATCAAGTCTTCAAAATCGGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	960
DB	1800	ATCAAGTCTTCAAAATCGGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	1859
QY	961	GGGCAACGCTGATTCACCTCAACAGAGATCTTCTTCAACACTGAAAGATAAGTCAAGTCA	1020
DB	1860	GGGCAACGCTGATTCACCTCAACAGAGATCTTCTTCAACACTGAAAGATAAGTCAAGTCA	1919
QY	1021	CCAGAGACGGAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAAATTTAGTCCAAAA	1080
DB	1920	CCAGAGACGGAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAAATTTAGTCCAAAA	1979
QY	1081	ACTTGAAGAGTACAGCAGACAGCTCTAGATTTACTGCAACAGTTCCTCCCTGAGCTGGA	1140
DB	1980	ACTTGAAGAGTACAGCAGACAGCTCTAGATTTACTGCAACAGTTCCTCCCTGAGCTGGA	2039
QY	1141	AAAGTTTCTTGCTGCTTACAGAGCTGGAACAACTGCCAATGTCTCTACAGGATGCTAC	1200
DB	2040	AAAGTTTCTTGCTGCTTACAGAGCTGGAACAACTGCCAATGTCTCTACAGGATGCTAC	2099
QY	1201	CCGTGAAGAAAGCTCTTGAAGACTCCAGAGGAGTAAAGAGCTGATGAACCAATGGCA	1260
DB	2100	CCGTGAAGAAAGCTCTTGAAGACTCCAGAGGAGTAAAGAGCTGATGAACCAATGGCA	2159
QY	1261	AGACCTCCAGAGTGAATTTGAAGCTCACACAGATGTTTATCACAACTCTGATGAAACAG	1320
DB	2160	AGACCTCCAGAGTGAATTTGAAGCTCACACAGATGTTTATCACAACTCTGATGAAACAG	2219
QY	1321	CAAAAAATCTCTGAGATCTCTGGAAGGTTCCGATGATGTCAGTCTCTTACAAAGAGCTTT	1380
DB	2220	CAAAAAATCTCTGAGATCTCTGGAAGGTTCCGATGATGTCAGTCTCTTACAAAGAGCTTT	2279
QY	1381	GGATAACATGAATCTCAAGTGGAGTGAATTCGGAAGGTTCCGATGATGTCAGTCTCTT	1440
DB	2280	GGATAACATGAATCTCAAGTGGAGTGAATTCGGAAGGTTCCGATGATGTCAGTCTCTT	2339
QY	1441	TTTGAAGCCAGTTCTGACCCAGTGGAAAGCTCTGCACTTTCTCTGAGGAACTTTCTGCT	1500
DB	2340	TTTGAAGCCAGTTCTGACCCAGTGGAAAGCTCTGCACTTTCTCTGAGGAACTTTCTGCT	2399
QY	1501	GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGACCTATTTGGAGGCGATTTCC	1560
DB	2400	GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGACCTATTTGGAGGCGATTTCC	2459
QY	1561	AGCAGTTTCAGAGACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACATAAGA	1620
DB	2460	AGCAGTTTCAGAGACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACATAAGA	2519
QY	1621	ACCTGTAATCATGATCTTTGAGCTGTAGATATTTCTGACAGAGACGCTTTGGA	1680
DB	2520	ACCTGTAATCATGATCTTTGAGCTGTAGATATTTCTGACAGAGACGCTTTGGA	2579
QY	1681	AGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCTCTGAGGAGAGAGAGAGAGAG	1740
DB	2580	AGGACTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCTCTGAGGAGAGAGAGAGAGAG	2639
QY	1741	TGTCATCTGGCTTTTACGAAAGCAGGCTGAGAGGTCATATCTAGTGGGAAATTTGAA	1800
DB	2640	TGTCATCTGGCTTTTACGAAAGCAGGCTGAGAGGTCATATCTAGTGGGAAATTTGAA	2699

1801 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAGACTCCAGAACT 1860
Db CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAGACTCCAGAACT 2759
1861 TCAGAGGCCACCGATGAGCTGAGACCTCAAGCTGCGCCAGCTGAGTGCATCAAGGATC 1920
Db TCAGAGGCCACCGATGAGCTGAGACCTCAAGCTGCGCCAGCTGAGTGCATCAAGGATC 2819
1921 CTGGCAGCCCGTGGGGATCTCTCTCAATGACTCTCTCCCAAGATCAGCTCGAGAAAGTCAA 1980
Db CTGGCAGCCCGTGGGGATCTCTCTCAATGACTCTCTCCCAAGATCAGCTCGAGAAAGTCAA 2879
1981 GGCACCTCGAGGAGAAATTCGGCTCTGAAAAGAACGTGAGCCAGCTCAATGACTTGC 2040
Db GGCACCTCGAGGAGAAATTCGGCTCTGAAAAGAACGTGAGCCAGCTCAATGACTTGC 2939
2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACGTTATACCTCAGCACCTCTGGAAGA 2100
Db TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCAACGTTATACCTCAGCACCTCTGGAAGA 2999
2101 C 2101
3000 C 3000
RESULT 2
AD37255
ID AD37255 standard; DNA; 5149 BP.
XX
AC AAD37255;
XX
JT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
DS Chimeric - Homo sapiens.
DS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13677.
XX
XX 28-APR-2000; 2000US-200777P.
XX
XX (XIAO/) XIAO X.
XX
XX Xiao X;
XX
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -
XX
XX Example 1; Page 57-59; 71pp; English.
XX
PS The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV

vector plasmid construct containing human dystrophin minigenes, a muscle
creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 5149 BP; 1489 A; 1236 C; 1269 G; 1155 T; 0 other;
Query Match 100.0%; Score 2101; DB 24; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGCTATGCCCTACACACAGAGCTGCTTATGTCAACACCTCTGACCTACACGAGCCCAATT 60
Db GAGCTATGCCCTACACACAGAGCTGCTTATGTCAACACCTCTGACCTACACGAGCCCAATT 1716
QY 61 TCCTTCACAGCATTTGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTCATTTGAGAG 120
Db TCCTTCACAGCATTTGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTCATTTGAGAG 1776
QY 121 TGAAGTAAACCTGACCGTTTATCAACAGCTTTTGAAGAGAGTATTATCGTGGCTTCTTTC 180
Db TGAAGTAAACCTGACCGTTTATCAACAGCTTTTGAAGAGAGTATTATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGTGAAAGA 240
Db TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTTAATGATGTGGAAGTGTGAAAGA 1896
QY 241 CCAGTTTCATCTACTGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 300
Db CCAGTTTCATCTACTGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 1956
QY 301 TAATATTCTCAATTTGGGAAGTAAAGCTGATTGGAACAGGAGAAATTTATCAGAAGATGAAGA 360
Db TAATATTCTCAATTTGGGAAGTAAAGCTGATTGGAACAGGAGAAATTTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGCGAATGCTTCAGGTAGC 420
Db AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGCGAATGCTTCAGGTAGC 2076
QY 421 TAGCATGCAAAAACAAAGCAATTTACATAGAGTTTTAATGATCTCCAGAAATCAGAAACT 480
Db TAGCATGCAAAAACAAAGCAATTTACATAGAGTTTTAATGATCTCCAGAAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACAAAGGAAATGAGGAGAA 540
Db GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAACAAAGGAAATGAGGAGAA 2196
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTACAACTAAGTGGTTC 600
Db GCCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTACAACTAAGTGGTTC 2256
QY 601 AGAAGATCTAGAACAAGAACCAAGTCAAGGTCAATTTCTCCTCCTCAGATGCTGGTAGT 660
Db AGAAGATCTAGAACAAGAACCAAGTCAAGGTCAATTTCTCCTCCTCAGATGCTGGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCAAGCACTGCTGCTTTGGAGAACAACTTAAGTATTGGG 720
Db TGATGAATCTAGTGGAGATCAAGCACTGCTGCTTTGGAGAACAACTTAAGTATTGGG 2376
QY 721 AGATCGATGGCAAAACATCTGTAGATGAGCAGAGAACCGCTGGGTCTTTTACAAGACAT 780
Db AGATCGATGGCAAAACATCTGTAGATGAGCAGAGAACCGCTGGGTCTTTTACAAGACAT 2436
QY 781 CCTTCTCAATGGCAACAGCTCTTACTGAAGAACAGTGGCTTTTGTAGTGCATGGCTTTCAGA 840
Db CCTTCTCAATGGCAACAGCTCTTACTGAAGAACAGTGGCTTTTGTAGTGCATGGCTTTCAGA 2496
QY 841 AAAAGAGATGCACTGAGCAAGATTTACACAACTGGCTTTTAAAGATCAAAATGTT 900
Db AAAAGAGATGCACTGAGCAAGATTTACACAACTGGCTTTTAAAGATCAAAATGTT 2556
QY 901 ATCAAGTCTTCAAAACCTGCGCGTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
Db ATCAAGTCTTCAAAACCTGCGCGTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 2616

QY 961 GGGCAAACTGATTCACCTCAACAGAGATCTCTTCAACACCTCAAGAAATAGTCAGTGAC 1020
Db 2617 GGGCAAACTGATTCACCTCAACAGAGATCTCTTCAACACCTCAAGAAATAGTCAGTGAC 2676
QY 1021 CCAGAACAGGAGCATGGCTGGATACCTTTGGCCGGTGTGGGATAATTTAGTCCAAA 1080
Db 2677 CCAGAACAGGAGCATGGCTGGATACCTTTGGCCGGTGTGGGATAATTTAGTCCAAA 2736
QY 1081 ACTTGAAGAGTACAGACAGACTCATAGATCTCAACAGTCCACAGTCCCTGGACCTGGA 1140
Db 2737 ACTTGAAGAGTACAGACAGACTCATAGATCTCAACAGTCCACAGTCCCTGGACCTGGA 2796
QY 1141 AAAGTTCTTCCTCGCTTACAGAGCTGAACACCTGCCAATGTCTACAGGATCTAC 1200
Db 2797 AAAGTTCTTCCTCGCTTACAGAGCTGAACACCTGCCAATGTCTACAGGATCTAC 2856
QY 1201 CCGTAAGAAAGGCTCTAGAGACTCAAGGAGTAAAGAGTGAATGAACAAATGGCA 1260
Db 2857 CCGTAAGAAAGGCTCTAGAGACTCAAGGAGTAAAGAGTGAATGAACAAATGGCA 2916
QY 1261 AGACCTCCAGGTCAAAATGAAGCTCACACAGATGTTTATCAACACCTGGATGAACACAG 1320
Db 2917 AGACCTCCAGGTCAAAATGAAGCTCACACAGATGTTTATCAACACCTGGATGAACACAG 2976
QY 1321 CCAAAAATCTGAGATCCCTGGAGGTTCCGATGTCAGTCTGTTCACAAAGAGTTT 1380
Db 2977 CCAAAAATCTGAGATCCCTGGAGGTTCCGATGTCAGTCTGTTCACAAAGAGTTT 3036
QY 1381 GGATAACATGAATCAAGTGGAGTGAATTCGGAAGAGTCTCTCAACATTAAGTCCCA 1440
Db 3037 GGATAACATGAATCAAGTGGAGTGAATTCGGAAGAGTCTCTCAACATTAAGTCCCA 3096
QY 1441 TTTGGAAGCCAGTTCTGACAGTGGAGGCTGCTGCACTTTCTCTGAGGAACTTCTGGT 1500
Db 3097 TTTGGAAGCCAGTTCTGACAGTGGAGGCTGCTGCACTTTCTCTGAGGAACTTCTGGT 3156
QY 1501 GTGGCTACAGTGAAGATGATGAATTAAGCCGAGGAGCACTATTGGAGGCACTTTC 1560
Db 3157 GTGGCTACAGTGAAGATGATGAATTAAGCCGAGGAGCACTATTGGAGGCACTTTC 3216
QY 1561 AGCAGTTCCAGAGCAGAACGATGATACATAGGCGCTTCAAGAGGGAATTTGAAACTAAAGA 1620
Db 3217 AGCAGTTCCAGAGCAGAACGATGATACATAGGCGCTTCAAGAGGGAATTTGAAACTAAAGA 3276
QY 1621 ACCTGTATCATGATCTCTGAGACTGTACGATATTTCTGACAGAGAGCTTTGGA 1680
Db 3277 ACCTGTATCATGATCTCTGAGACTGTACGATATTTCTGACAGAGAGCTTTGGA 3336
QY 1681 AGGACTAGAGAACTCTACAGGAGCCAGAGAGTGGCTCTCTGAGGAGAGAGCCAGAA 1740
Db 3337 AGGACTAGAGAACTCTACAGGAGCCAGAGAGTGGCTCTCTGAGGAGAGAGCCAGAA 3396
QY 1741 TGTCTCTGGCTTCTACGAAAGAGGCTGAGAGGTCAATCTAGTGGGAAATTTGAA 1800
Db 3397 TGTCTCTGGCTTCTACGAAAGAGGCTGAGAGGTCAATCTAGTGGGAAATTTGAA 3456
QY 1801 CTGCACTCCGCTGCTGACAGAGAAAATAGATGAGACCTTCAAGAGCTTCCAGAGACT 1860
Db 3457 CTGCACTCCGCTGCTGACAGAGAAAATAGATGAGACCTTCAAGAGCTTCCAGAGACT 3516
QY 1861 TCAAGAGGCCAGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGTGAATCAAGGATC 1920
Db 3517 TCAAGAGGCCAGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGTGAATCAAGGATC 3576
QY 1921 CTGGAGCCGCTGGGGATCTCTCATTTGACTCTCTCAAGATCACTTCGAGAAAGTCAA 1980
Db 3577 CTGGAGCCGCTGGGGATCTCTCATTTGACTCTCTCAAGATCACTTCGAGAAAGTCAA 3636
QY 1981 GGCACCTTCGAGAGAAATTCGCTCTGAAGAGAACTGAGCCACAGTCAATGACCTTGC 2040
Db 3637 GGCACCTTCGAGAGAAATTCGCTCTGAAGAGAACTGAGCCACAGTCAATGACCTTGC 3696
QY 2041 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACGGTATTAACCTTCAGCACTCTCTGGAAGA 2100

Db 3697 TCGCCAGCTTACCACCTTTGGGCAATTCAGCTCTACCGTATATACCTCAGCACTCTGGAAGA 3756
QY 2101 C 2101
Db 3757 C 3757
RESULT 3
AAD37234
ID AAD37234 standard; DNA; 3999 BP.
XX AAD37234;
AC
DT 21-AUG-2002 (first entry)
XX
XX Human dystrophin minigene delta3990.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
XX WO200183695-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13677.
XX
XX 28-APR-2000; 2000US-200777P.
XX
XX (XIAO/) XIAO X.
XX
XX Xiao X;
XX
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
XX Example 1; Page 46-47; 71pp; English.
PS
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2, 8059-10227 (rods R22, R23 and R24, hinge
CC H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
XX Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 other;
SQ
Query Match 75.2%; Score 1579.8; DB 24; Length 3999;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTCTACACCTCTGACCTCCTACCGAGCCCAATT 60
Db 900 GAGCTATGCTTACACACAGGCTGCTTATGTCTACACCTCTGACCTCCTACCGAGCCCAATT 959
QY 61 TCCTTCACAGCAATTTGGAGGCTCTTGACAGCAAGTCATTTGGCACTTCATTGATGAGAG 120
Db 960 TCCTTCACAGCAATTTGGAGGCTCTTGACAGCAAGTCATTTGGCACTTCATTGATGAGAG 1019

121 TGAAGTAAACCTCGACCGTTATCAACAGCTTTAGAGAGATTAATCGTGGCTTCTTC 180
120 TGAAGTAAACCTCGACCGTTATCAACAGCTTTAGAGAGATTAATCGTGGCTTCTTC 1079
181 TGTGAGGACACATTCGAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1080 TGTGAGGACACATTCGAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
241 CCAAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300
1140 CCAAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1199
301 TAATATTTCAATTTGGAGAGTAAAGTGAATGGAACAGGAAATTAATCAAGAGATGAAGA 360
1200 TAATATTTCAATTTGGAGAGTAAAGTGAATGGAACAGGAAATTAATCAAGAGATGAAGA 1259
361 AACTGAAGTACAGAGCAGATGATCTCTTAATTCAGATGGGATGCTCAGGGTAC 420
1260 AACTGAAGTACAGAGCAGATGATCTCTTAATTCAGATGGGATGCTCAGGGTAC 1319
421 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAATGGATCTCCAGAAATCAGAACT 480
1320 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAATGGATCTCCAGAAATCAGAACT 1379
481 GAAAGAGTTGAATGACTGGCTTAACAAACAGAGAAAGACAGGAAATGAGAGAGA 540
1380 GAAAGAGTTGAATGACTGGCTTAACAAACAGAGAAAGACAGGAAATGAGAGAGA 1439
541 GCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAAGGTGCTTCA 600
1440 GCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAAGGTGCTTCA 1499
601 AGAAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTCAGTCAATGAGTGTAGT 660
1500 AGAAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTCAGTCAATGAGTGTAGT 1559
661 TGATGAATCTAGTGGAGATCAGCAACCTGCTTTGGAAGAACAACTTAAGGTATTGG 720
1560 TGATGAATCTAGTGGAGATCAGCAACCTGCTTTGGAAGAACAACTTAAGGTATTGG 1619
721 AGATCGATGGCAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTCAAGACAT 780
1620 AGATCGATGGCAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTCAAGACCA 1679
781 CCTTCTCAATGGCAAGCTCTTACTGAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
1680 GCCTGAC-----CTAGCTCTGAGTACCACTATTGGAGCTCTCCTACTCA 1727
841 AAAAGAGATGCAAGTGAACAGATTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
1728 GACTGTGTACTCTGCTGACACA-----ACCTGTGTTTACTAAGGAACTGCCATCT- 1777
901 ATCAAGTCTTCAAAAACCTGGCCGTTTTAAAGCGGATCTAGAAAGAAAGCAATCCAT 960
1778----- 1777
961 GGGCAAACTGTATTCACTCAAAACAGATCTCTTTCAACACTGAGAAATAGTCAGTGAC 1020
1778-----CCAAACTAGAAATGCCATCTCTCTTGATGTGGAG----- 1812
1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTTAGTCCAAAA 1080
1813----- 1812
1081 ACTTGAAGAGTACAGACAGATCTCATAGATTACTGCAACAGTTCCCTGGACCTGGA 1140
1813-----GTACTACTCATAGATTACTGCAACAGTTCCCTGGACCTGGA 1856
1141 AAAGTTCTTGCTGGCTTACAGAGCTGAAACAACTGCCAATCTCTACAGGATGCTAC 1200
1857 AAAGTTCTTGCTGGCTTACAGAGCTGAAACAACTGCCAATCTCTACAGGATGCTAC 1916
1201 CCGTAAGGAAGGCTCTTAGAGACTCCAGGGAGTAAAGAGCTGATGAACAATGCCA 1260

1917 CCGTAAGGAAGGCTCTTAGAGACTCCAGGGAGTAAAGAGCTGATGAACAATGCCA 1976
1261 AGACCTCAAGGTGAAATTCGAAGCTCACAGATGTTTATCAACAACCTCGATGAACAACAG 1320
1977 AGACCTCAAGGTGAAATTCGAAGCTCACAGATGTTTATCAACAACCTCGATGAACAACAG 2036
1321 CCAAAAAATCCTGAGATCCCTCGAGAGGTTCCGATGATGAGTCTCTGTACAAAGACGTTT 1380
2037 CCAAAAAATCCTGAGATCCCTCGAGAGGTTCCGATGATGAGTCTCTGTACAAAGACGTTT 2096
1381 GGATAACATCAACTTCAAGTGGAGTGAACCTCGAAAAAGTCTCTCAACATTAGGTCCCA 1440
2097 GGATAACATCAACTTCAAGTGGAGTGAACCTCGAAAAAGTCTCTCAACATTAGGTCCCA 2156
1441 TTTGGAAGCCAGTTTCTGACCAAGTGAAGGTTCTGCACCTTTCTCTGAGGAACTTCTGGT 1500
2157 TTTGGAAGCCAGTTTCTGACCAAGTGAAGGTTCTGCACCTTTCTCTGAGGAACTTCTGGT 2216
1501 GTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGCACCTATTGAGGCGGACTTTCC 1560
2217 GTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGCACCTATTGAGGCGGACTTTCC 2276
1561 AGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAGA 1620
2277 AGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAGA 2336
1621 ACCTGTAACTAGTACTCTTGTAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 1680
2337 ACCTGTAACTAGTACTCTTGTAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 2396
1681 AGGACTAGAGAACTCTACCAAGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAA 1740
2397 AGGACTAGAGAACTCTACCAAGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAA 2456
1741 TGTCACTCGGCTTCTACGAAAGAGGCTGAGAGGTCATCTAGTGGGAAAAATTTGAA 1800
2457 TGTCACTCGGCTTCTACGAAAGAGGCTGAGAGGTCATCTAGTGGGAAAAATTTGAA 2516
1801 CCTGACCTCCGCTGACGTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCACAGGAAT 1860
2517 CCTGACCTCCGCTGACGTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCACAGGAAT 2576
1861 TCAAGAGCCCAAGATGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGTGTATCAAGGATC 1920
2577 TCAAGAGCCCAAGATGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGTGTATCAAGGATC 2636
1921 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAA 1980
2637 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAA 2696
1981 GGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAAAGAACGTGAGCCACGTCATACCTTGC 2040
2697 GGCACCTTCGAGGAGAAATTTGGCCCTCTGAAAAAGAACGTGAGCCACGTCATACCTTGC 2756
2041 TGCACGCTTACCACTTTGGGCAATTCAGTCTCACCTGATATACCTCAGCCTCTGGAGA 2100
2757 TGCACGCTTACCACTTTGGGCAATTCAGTCTCACCTGATATACCTCAGCCTCTGGAGA 2816
2101 C 2101
2817 C 2817

RESULT 4

AAD37256

ID AAD37256 standard; DNA; 4966 BP.

XX AAD37256;

XX AAD37256;

XX 21-AUG-2002 (first entry)

XX Adeno-associated virus vector plasmid, AAV-MCK-delta3990.

DE

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX WO200183695-A2.
XX 08-NOV-2001.
PD 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
XX Xiao X;
XX WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX Example 1; Page 59-60; 71pp; English.
PS The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;
SQ Query Match 75.2%; Score 1579.8; DB 24; Length 4966;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;
QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAAT 60
DB 1657 GAGCTATGCTTACACACAGGCTGCTTATGTACACACCTCTGACCCCTACACGGAGCCCAAT 1716
QY 61 TCCTTCACAGCATTTGGAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
DB 1717 TCCTTCACAGCATTTGGAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTCATTGATGGAGAG 1776
QY 121 TGAAGTAACTGGACCTGATATCAACAGCTTTTGAAGAGAGTATATCGTGGCTTCTTTC 180
DB 1777 TGAAGTAACTGGACCTGATATCAACAGCTTTTGAAGAGAGTATATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTGCAAGACAAGGAGAGATTTCTAATGATGTGGAATGGTGAAGA 240
DB 1837 TGCTGAGGACACATTGCAAGACAAGGAGAGATTTCTAATGATGTGGAATGGTGAAGA 1896
QY 241 CCAGTTTCTACTCATGAGGGGTACATGATGATTTTGACGCCCATCAGGCCGGGTGG 300
DB 1897 CCAGTTTCTACTCATGAGGGGTACATGATGATTTTGACGCCCATCAGGCCGGGTGG 1956
QY 301 TAAATATCTTACAAATTTGGGAAGTAACTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 360
DB 1957 TAAATATCTTACAAATTTGGGAAGTAACTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGATGGGATGCCTCAGGGTAGC 420

DB 2017 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATCCCTCAGGGTAGC 2076
QY 421 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTATATGGATCTCCAGAATCAGAAACT 480
DB 2077 TAGCATGGAACAAACAAAGCAATTTACATAGAGTTTATATGGATCTCCAGAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGACTGCTGCTAACAAAAACAGAGAAACAGAAAGAAATGGAGGAAGA 540
DB 2137 GAAAGAGTTGAATGACTGCTGCTAACAAAAACAGAGAAACAGAAAGAAATGGAGGAAGA 2196
QY 541 GCCTCTTGACCTGATCTTTGAAGACCTAAAGCCCAAGTACAAACCAATAAGGCTCTTCA 600
DB 2197 GCCTCTTGACCTGATCTTTGAAGACCTAAAGCCCAAGTACAAACCAATAAGGCTCTTCA 2256
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAAGTCAATCTCTCACTCACTGCTGGTGTAGT 660
DB 2257 AGAAGATCTAGAACAAAGAACAAAGTCAAGTCAATCTCTCACTCACTGCTGGTGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTCTTTGGAAGAACAACTTAAAGGTATTGGG 720
DB 2317 TGATGAATCTAGTGGAGATCAGCAACTGCTCTTTGGAAGAACAACTTAAAGGTATTGGG 2376
QY 721 AGATCGATGGCAAAACATCTGTAGATGGAACAAGACCTGCTGGTCTTTTACAGACAT 780
DB 2377 AGATCGATGGCAAAACATCTGTAGATGGAACAAGACCTGCTGGTCTTTTACAGACCA 2436
QY 781 CCTTCTCAATGGCAAGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGA 840
DB 2437 GCCTGAC-----CTAGCTCTGTGACCTGACCATATTGGAGCCTCTCTTACTCA 2484
QY 841 AAAAGAGATGTCAGTGAACAAAGATTCACACAACTGGCTTTAAAGATCAAAATGAATGTT 900
DB 2485 GACTGTACTCTGTGTACACA-----ACCTGTGGTTACTTAAGGAAACTGCCATCT- 2534
QY 901 ATCAAGTCTTCAAAAACCTGCCCTTTTAAAGCGGATCTAGAAAGCAAAAGCAATCCAT 960
DB 2535 ----- 2534
QY 961 GGGCAAACTGTATTACTCAAAACAAGATCTTTTCAACACTGAAGAATAAGTCAGTAC 1020
DB 2535 -----CCTGAGTGAATGCTTCTCTGATGTTGGAG----- 2569
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCGGTGTGGGATAATTAGTCCAAAA 1080
DB 2570 ----- 2569
QY 1081 ACTTGAAGAGATGACACAGACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGA 1140
DB 2570 -----GTACCTACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGA 2613
QY 1141 AAAAGTTTCTGCTGGCTTACAGAAAGCTGAAAACAACTGCCAATGTCTCAGAGTGTAC 1200
DB 2614 AAAAGTTTCTGCTGGCTTACAGAAAGCTGAAAACAACTGCCAATGTCTCAGAGTGTAC 2673
QY 1201 CGGTGAAGAAAGGCTCTCTAGAGACTCCAGGGAGTAAAGAGCTGATGAAACAAATGGCA 1260
DB 2674 CGGTGAAGAAAGGCTCTCTAGAGACTCCAGGGAGTAAAGAGCTGATGAAACAAATGGCA 2733
QY 1261 AGACCTTCAAGGTGAATTTGAAGCTCAACAGATGTTTATCACAACCTGGATGAACAG 1320
DB 2734 AGACCTTCAAGGTGAATTTGAAGCTCAACAGATGTTTATCACAACCTGGATGAACAG 2793
QY 1321 CAAAATAATCTGAGATCCCTGGAGGTTCCGATGATGCTGCTGTTTACAAAACGCTTT 1380
DB 2794 CAAAATAATCTGAGATCCCTGGAGGTTCCGATGATGCTGCTGTTTACAAAACGCTTT 2853
QY 1381 GGATAACATGAATTTCAAGTGGAGTGAATCTTGGAAAAAGTCTCTCAACATTAGTCCCA 1440
DB 2854 GGATAACATGAATTTCAAGTGGAGTGAATCTTGGAAAAAGTCTCTCAACATTAGTCCCA 2913
QY 1441 TTTGGAGCCAGTCTGTGACCAAGCTGTGACCTTTCTCTGACAGGAATCTTCTGTT 1500

2914 TTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACCTTCTGCT 2973
2Y GTGGCTACAGCTGAAGATGATGATTAAGCCGCGACCGACCTATTTGGGCGACTTTCC 1560
DB GTGGCTACAGCTGAAGATGATGATTAAGCCGCGACCGACCTATTTGGGCGACTTTCC 3033
2Y AGCAGTTTCAAGACAGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB AGCAGTTTCAAGACAGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3093
2Y ACTGTAAATCATGAGTACTCTTGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1680
DB ACTGTAAATCATGAGTACTCTTGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 3153
2Y AGGACTAGAGAACTCTACAGAGCCAGAGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1740
DB AGGACTAGAGAACTCTACAGAGCCAGAGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 3213
2Y TGTCACTTCGCTCTTACGAAAGCAGGCTGAGGAGTCAATATCTGAGTGGGAAATTTGAA 1800
DB TGTCACTTCGCTCTTACGAAAGCAGGCTGAGGAGTCAATATCTGAGTGGGAAATTTGAA 3273
2Y CCTGCACTTCGCTCTTACGAAAGCAGGCTGAGGAGTCAATATCTGAGTGGGAAATTTGAA 1860
DB CCTGCACTTCGCTCTTACGAAAGCAGGCTGAGGAGTCAATATCTGAGTGGGAAATTTGAA 3333
2Y TCAAGAGCCACAGGATGAGTGGAGCTCAAGCTGCGCCAGAGTGGAGTGGAGTGGAGTGGAGT 1920
DB TCAAGAGCCACAGGATGAGTGGAGCTCAAGCTGCGCCAGAGTGGAGTGGAGTGGAGTGGAGT 3393
2Y CTGGCAGCCGCTGGGCGATCTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCT 1980
DB CTGGCAGCCGCTGGGCGATCTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCT 3453
2Y GGCACCTTCGAGGAGAAATTTGGCTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTG 2040
DB GGCACCTTCGAGGAGAAATTTGGCTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTG 3513
2Y TCGCCAGCTTACACATTTGGGCGATGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCT 2100
DB TCGCCAGCTTACACATTTGGGCGATGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCT 3573
2Y 2101 C 2101
DB 3574 C 3574

RESULT 5

AAD37262
ID AAD37262 standard; DNA; 4990 BP.
XX AAD37262;
AC AAD37262;
DT 21-AUG-2002 (first entry)
XX
XX Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
XX WO200183695-A2.
PN 08-NOV-2001.
PD 27-APR-2001; 2001WO-US13677.
XX
XX 28-APR-2000; 2000US-20077P.
PR
XX

(XIAO/) XIAO X.

Xiao X;

WPI; 2002-049342/06.

New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -

Example 1; Page 67-68; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a cytomagalovirus (CMV) promoter and a small polyA signal sequence.

Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 other;

Query Match 75.2%; Score 1579.8; DB 24; Length 4990;

Best Local Similarity 87.9%; Pred. No. 0;

Matches 1846; Conservative 0; Mismatches 72; Indels 183; Gaps 4;

QY 1 GAGCTATGCTTACACACAGAGTCTGCTATGTACACACCTCTGACCTACACGAGCCCAT 60
DB 1681 GAGCTATGCTTACACACAGAGTCTGCTATGTACACACCTCTGACCTACACGAGCCCAT 1740
QY 61 TCCTTCACAGCATTTGGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTCATTGATGGAGAG 120
DB 1741 TCCTTCACAGCATTTGGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTCATTGATGGAGAG 1800
QY 121 TGAAGTAACTGACCGTTATCAACAGCTTTTGAAGAAGTATTATCGTGGCTCTTTTC 180
DB 1801 TGAAGTAACTGACCGTTATCAACAGCTTTTGAAGAAGTATTATCGTGGCTCTTTTC 1860
QY 181 TCCTGAGGACACATTTGGAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA 240
DB 1861 TCCTGAGGACACATTTGGAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA 1920
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCCCGGGTTGG 300
DB 1921 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCCCGGGTTGG 1980
QY 301 TAATATTTCTACAATTTGGAAGTAAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 360
DB 1981 TAATATTTCTACAATTTGGAAGTAAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 2040
QY 361 AACTCAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGTTAGC 420
DB 2041 AACTCAAGTACAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGTTAGC 2100
QY 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTTAAATGATCTCCAGATCAGAAACT 480
DB 2101 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTTAAATGATCTCCAGATCAGAAACT 2160
QY 481 GAAAGAGTTGAATGATCTGGCTAAACAAAACAGAGAAGAAACAAGGAAATTTGGAAGGAAGA 540
DB 2161 GAAAGAGTTGAATGATCTGGCTAAACAAAACAGAGAAGAAACAAGGAAATTTGGAAGGAAGA 2220
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACACATAGGTGCTTCA 600
DB 2221 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAACACATAGGTGCTTCA 2280
QY 601 AGAAGATCTTAGAACAAGAACCAAGTCAGGGTCAATTTCTCTCACTCACATGTTGGTGTAGT 660

Db 2281 AGAAGATCTAGAAACAGAAACAGTCCGGTCAATCTCTCACTACATGCTGGTGGTAGT 2340
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAAGAAACAATTAAGATTGGG 720
Db 2341 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAAGAAACAATTAAGATTGGG 2400
QY 721 AGATCGATGGGCAACACTGTAGATGACAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 2401 AGATCGATGGGCAACACTGTAGATGACAGACCGCTGGGTTCTTTTACAAGACCA 2460
QY 781 CCTTCTCAAAATGGCAACGCTTTACTGAAGACAGTGCCTTTTATAGTCATGCTTTTCA 840
Db 2461 GCTGAC-----CTAGCTCCTGGACTGACCACTATTGGAGCCTCTCCTACTCA 2508
QY 841 AAAGAGATGAGTCAACAGCAATTCACACACTGGCTTTAAAGATCAAAATGAATGTT 900
Db 2509 GACTGTACTCTGGTGACACA-----ACCTGGTGTACTTAAGGAACATGCCATCT- 2558
QY 901 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 2559 ----- 2558
QY 961 GGGCAAACTGTATTCACTCAAAACAGATCTCTTTCAACACTGACAGTAATAGTCAGTCAC 1020
Db 2559 -----CCAAAACTAGAAATGCCATCTTCTTTGATGTTGGAG----- 2593
QY 1021 CCAGAGACGGAAGCATGCTGGATAACTTTTCCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 2594 ----- 2593
QY 1081 ACTTGAAAGAGTACAGACACAGCTCATAGATTACTGCAACAGTTCCCGCTGGAGCTGGA 1140
Db 2594 -----GTACTACTCATAGATTACTGCAACAGTTCCCGCTGGAGCTGGA 2637
QY 1141 AAAGTTTCTGCTGCTTACAGAGCTCAACAACTGCCAATGCTTACAGAGATGCTAC 1200
Db 2638 AAAGTTTCTGCTGCTTACAGAGCTCAACAACTGCCAATGCTTACAGAGATGCTAC 2697
QY 1201 CCCTAAGGAAAGGCTCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAACAATGGCA 1260
Db 2698 CCCTAAGGAAAGGCTCTAGAGACTCCAAAGGAGTAAAGAGCTGATGAACAATGGCA 2757
QY 1261 AGACTCCAAAGTGAATTTGAAGCTCACACAGATGTTTATCAACCTGGATGAAGAACAG 1320
Db 2758 AGACTCCAAAGTGAATTTGAAGCTCACACAGATGTTTATCAACCTGGATGAAGAACAG 2817
QY 1321 CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCCTGTTTCAAGAGACGTTT 1380
Db 2818 CCAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGAGTCCCTGTTTCAAGAGACGTTT 2877
QY 1381 GGATAACATGAATTTCAAGTGGAGTGAACCTTGGAAAAAGTCTCTCAACATTAGTCCCA 1440
Db 2878 GGATAACATGAATTTCAAGTGGAGTGAACCTTGGAAAAAGTCTCTCAACATTAGTCCCA 2937
QY 1441 TTTGGAGGACGATCTTGACAGTGGAGGCTGTGACCTTCTCTCAGAGAACTTCTGTT 1500
Db 2938 TTTGGAGGACGATCTTGACAGTGGAGGCTGTGACCTTCTCTCAGAGAACTTCTGTT 2997
QY 1501 GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCACTTTCC 1560
Db 2998 GTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGCACTTTCC 3057
QY 1561 AGCATTTCAAGACAGAACCTGTACATAGGGCTTCAAGAGGGAATTTGAAGAACTAAGA 1620
Db 3058 AGCATTTCAAGACAGAACCTGTACATAGGGCTTCAAGAGGGAATTTGAAGAACTAAGA 3117
QY 1621 ACCTGTAAATCATGAGTACTCTTGAGCTGTACGAATATTTCTGACAGACAGCCTTTTGA 1680
Db 3118 ACCTGTAAATCATGAGTACTCTTGAGCTGTACGAATATTTCTGACAGACAGCCTTTTGA 3177
QY 1681 AGGACTAGAGAACTTACCGAGGACCCAGAGAGTCCCTCTGAGGAGAGCCAGAA 1740

Db 3178 AGGACTAGAGAAACTTACTACGAGGCCCCAGAGAGCTGCCTCTCTGAGGAGAGAGCCAGAA 3237
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAAAATTGAA 1800
Db 3238 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAAAATTGAA 3297
QY 1801 CTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGAACT 1860
Db 3298 CTTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGAACT 3357
QY 1861 TCAAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 1920
Db 3358 TCAAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 3417
QY 1921 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCTGAGAAAGTCAA 1980
Db 3418 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCTGAGAAAGTCAA 3477
QY 1981 GGCACCTCGAGGAGAAATTTGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 2040
Db 3478 GGCACCTCGAGGAGAAATTTGCCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC 3537
QY 2041 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 2100
Db 3538 TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGA 3597
QY 2101 C 2101
Db 3598 C 3598

RESULT 6
AAD37237 standard; DNA; 3858 BP.
XX AAD37237;
XX 21-AUG-2002 (first entry)
XX Human dystrophin minigene delta3849.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; EMD;
XX Becker muscular dystrophy; ds.
XX Homo sapiens.
XX WO2001:83695-A2.
XX 08-NOV-2001.
XX 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
XX Xiao X;
XX WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -
XX Example 1; Page 48-49; 71pp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)

CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R2, R3 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;

Query Match 68.7%; Score 1443; DB 24; Length 3858;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 177; Conservative 0; Mismatches 0; Indels 324; Gaps 1;

QY 1 GAGCTATGCCCTACACAGAGCTGCTTATGTCACACCTCTGACCTACACGAGCCCAT 60
DB 900 GAGCTATGCCCTACACAGAGCTGCTTATGTCACACCTCTGACCTACACGAGCCCAT 959

QY 61 TCCTTCACAGATTGGAAGCTCTGAAGACAAGTCAATTTGGCAGTTCAATGATGAGAG 120
DB 960 TCCTTCACAGATTGGAAGCTCTGAAGACAAGTCAATTTGGCAGTTCAATGATGAGAG 1019

QY 121 TGAAGTAACTGACCGTCTATCAACAGCTTTAGAGAGTATATCGTGGCTCTTTTC 180
DB 1020 TGAAGTAACTGACCGTCTATCAACAGCTTTAGAGAGTATATCGTGGCTCTTTTC 1079

QY 181 TGCTGAGGACACATTGCAAGCACAGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
DB 1080 TGCTGAGGACACATTGCAAGCACAGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1139

QY 241 CCAGTTTCATCTCATCAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 300
DB 1140 CCAGTTTCATCTCATCAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 1199

QY 301 TAATATTTCTCAATTTGGAAGTAGTGTGATTTGGAACAGAGAAATTTATCAGAGATGAAGA 360
DB 1200 TAATATTTCTCAATTTGGAAGTAGTGTGATTTGGAACAGAGAAATTTATCAGAGATGAAGA 1259

QY 361 AACTGAGTACAGAGAGATGAATCTCTTAATTTAAGTGGATGCTCAGGTTAGC 420
DB 1260 AACTGAGTACAGAGAGATGAATCTCTTAATTTAAGTGGATGCTCAGGTTAGC 1319

QY 421 TAGCATGGAACAAACAGCAATTTACATAGATTTTAAATGATCTCCAGAAATCAGAAACT 480
DB 1320 TAGCATGGAACAAACAGCAATTTACATAGATTTTAAATGATCTCCAGAAATCAGAAACT 1379

QY 481 GAAAGAGTTGAATCACTGGCTTAACAAAACAGAGAAAGAACAGGAAATGAGGAAGA 540
DB 1380 GAAAGAGTTGAATCACTGGCTTAACAAAACAGAGAAAGAACAGGAAATGAGGAAGA 1439

QY 541 GCCTCTTGGACCTGATCTTGAGAGCTTAAACGCCAAGTACACACATAGGTGCTTCA 600
DB 1440 GCCTCTTGGACCTGATCTTGAGAGCTTAAACGCCAAGTACACACATAGGTGCTTCA 1499

QY 601 AGAAGATCTAGAACAGAACAGTCAAGGTCAATTTCTCACTCACATGGTGGTGTAGT 660
DB 1500 AGAAGATCTAGAACAGAACAGTCAAGGTCAATTTCTCACTCACATGGTGGTGTAGT 1559

QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAACCAATTAAGTATTGGG 720
DB 1560 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAACCAATTAAGTATTGGG 1619

QY 721 AGATCGATGGGCAACACTCTGTAGATGACAGAGAGCGGTGGGTCTTTTACAAGACAT 780
DB 1620 AGATCGATGGGCAACACTCTGTAGATGACAGAGAGCGGTGGGTCTTTTACAAGACAT 1677

QY 781 CTTCTCAATTTGGCAACAGCTCTTAATGAGAACAGTGCCTTTTATGTCATGGCTTTTCAGA 840
DB 1678 CTTCTCAATTTGGCAACAGCTCTTAATGAGAACAGTGCCTTTTATGTCATGGCTTTTCAGA 1677

QY 841 AAAAGAGATGCAGTGAACAGATTTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900

DB 1678 ----- 1677
QY 901 ATCAAGTCTTCAAAAACCTGGCGGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
DB 1678 ----- 1677
QY 961 GGGCAAACTGTAATTCATCTCAAAACAGATCTTCTTCAACTGAGAAATAGTCAGTGAC 1020
DB 1678 ----- 1677
QY 1021 CCAGAGACGGAAGCATGCTGGATAACTTTGCCGGGTGTTGGGATAAATTTAGTCCAAA 1080
DB 1678 ----- 1677
QY 1081 ACTTGAAAAGAGTACAGCAGACAGATCTAGATTAATCTGCAACAGTTCCCTCTGGACCTTGA 1140
DB 1678 ----- ACTCATAGATTACTGCAACAGTCCCTCTGGACCTTGA 1715
QY 1141 AAAGTTTCTTCCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 1200
DB 1716 AAAGTTTCTTCCCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTAC 1775
QY 1201 CCCTAAGGAAAAGGCTCTCTAGAGACTCCAAAGGGAGTAAAAGAGCTGATGAAACAAATGGCA 1260
DB 1776 CCCTAAGGAAAAGGCTCTCTAGAGACTCCAAAGGGAGTAAAAGAGCTGATGAAACAAATGGCA 1835
QY 1261 AGACCTCCAAAGTGAATTTGAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1320
DB 1836 AGACCTCCAAAGTGAATTTGAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAG 1895
QY 1321 CCAAAAATCTCTGAGATCCCTGGAAGTTCCGATGATGCGAGTCTCTTACAAAGACGTTT 1380
DB 1896 CCAAAAATCTCTGAGATCCCTGGAAGTTCCGATGATGCGAGTCTCTTACAAAGACGTTT 1955
QY 1381 GGATTAACATGAATTTCAAGTGGAGTGAATTTGCGAAAAGTCTCTCAACATAGGTCCCA 1440
DB 1956 GGATTAACATGAATTTCAAGTGGAGTGAATTTGCGAAAAGTCTCTCAACATAGGTCCCA 2015
QY 1441 TTTGGAACCCAGTTCTGTACAGTGGAAAGCTCTGCACCTTCTCTCGAGGAATCTCTGCT 1500
DB 2016 TTTGGAACCCAGTTCTGTACAGTGGAAAGCTCTGCACCTTCTCTCGAGGAATCTCTGCT 2075
QY 1501 GTGGCTACAGCTGAAAAGATGATGAATTAAGCCGCGAGGACCTTAATTTGGAGGCGACTTCC 1560
DB 2076 GTGGCTACAGCTGAAAAGATGATGAATTAAGCCGCGAGGACCTTAATTTGGAGGCGACTTCC 2135
QY 1561 AGCAGTTTCAGAGCAGACAGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACCTAAGA 1620
DB 2136 AGCAGTTTCAGAGCAGACAGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACCTAAGA 2195
QY 1621 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACGAGCAGCCTTTTGA 1680
DB 2196 ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACGAGCAGCCTTTTGA 2255
QY 1681 AGGACTAGAGAACTCTACAGAGGCGCCAGAGCTGCTCTCTCAGGAGAGAGCCAGAA 1740
DB 2256 AGGACTAGAGAACTCTACAGAGGCGCCAGAGAGCTGCTCTCTCAGGAGAGAGCCAGAA 2315
QY 1741 TGTCACTCGGCTTTTACGAAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAATTTGAA 1800
DB 2316 TGTCACTCGGCTTTTACGAAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGAAAATTTGAA 2375
QY 1801 CCGTCACTCCGCTGATGCGAGAGAAAATATAGATGAGACCCCTTGAAGACTCCAGGACT 1860
DB 2376 CCGTCACTCCGCTGATGCGAGAGAAAATATAGATGAGACCCCTTGAAGACTCCAGGACT 2435
QY 1861 TCAAGAGCCACGGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 1920
DB 2436 TCAAGAGCCACGGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGGATC 2495
QY 1921 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACTCTCGAGAAAGTCAA 1980
DB 2496 CTGGCAGCCCGTGGCGATCTCTCATTTGACTCTCTCAAGATCACTCTCGAGAAAGTCAA 2555

QY	1981	GGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTCAGCCACGTCATATGACCTTGC	2040
Db	2556	GGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACGTCAGCCACGTCATATGACCTTGC	2615
QY	2041	TGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGGTATTAACCTCAGCACTCTGGAAGA	2100
Db	2616	TGCCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGGTATTAACCTCAGCACTCTGGAAGA	2675
QY	2101	C 2101	
Db	2676	C 2676	
RESULT 7			
ID	AAD37257		
AD	AAD37257	standard; DNA; 4825 BP.	
XX			
AC	AAD37257;		
DT	21-AUG-2002	(first entry)	
XX			
DE	Adeno-associated virus vector plasmid, AAV-MCK-delta3849.		
XX			
KW	Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;		
KW	adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;		
KW	Becker muscular dystrophy; ds.		
OS	Chimeric - Homo sapiens.		
OS	Chimeric - Unidentified.		
XX			
PN	WO200183695-A2.		
XX			
PD	08-NOV-2001.		
XX			
PF	27-APR-2001; 2001WO-US13677.		
XX			
PR	28-APR-2000; 2000US-200777P.		
XX			
PA	(XIAO/) XIAO X.		
PI	Xiao X;		
XX			
DR	WPI; 2002-049342/06.		
XX			
PT	New dystrophin minigene for treating Duchenne or Becker muscular		
PT	dystrophy comprises an N-terminal domain or modified N-terminal domain,		
PT	rod repeats, H1 and H4 domains and a cysteine rich domain of a		
PT	dystrophin gene		
XX			
PS	Example 1; Page 61-62; 71pp; English.		
XX			
CC	The present invention relates to an isolated nucleotide sequence encoding		
CC	a dystrophin minigene. The minigene comprises N-terminal or modified		
CC	N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4		
CC	domains and cysteine-rich domains of dystrophin or utrophin genes. The		
CC	invention also relates to a recombinant adeno-associated virus (AAV)		
CC	comprising dystrophin minigene operably linked to an expression control		
CC	element. The dystrophin minigene in operable linkage with an expression		
CC	control element, in a recombinant adeno-associated virus or retrovirus is		
CC	useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular		
CC	dystrophy (BMD) in a mammalian subject. The present sequence is AAV		
CC	vector plasmid construct containing human dystrophin minigenes, a muscle		
CC	creatine kinase (MCK) promoter and a small polyA signal sequence.		
XX			
SQ	Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 other;		
XX			
Query Match	68.7%; Score 1443; DB 24; Length 4825;		
Best Local Similarity	84.6%; Pred. No. 0;		
Matches 1777; Conservative	0; Mismatches 0; Indels 324; Gaps 1;		
QY	1	GAGCTAGCTTACACAGAGCTGCTTATGTCACACCTCTGACCTTACAGGAGCCATT	60
XX			

QY	1141	AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGATGCTAC	1200
Db	2473	AAAGTTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTTACAGATGCTAC	2532
QY	1201	CCGTAAAGGAAGCTCCTAGAGACTCCAGGGAGTAAAGAGCTCATGAACAATGGCA	1260
Db	2533	CCGTAAAGGAAGCTCCTAGAGACTCCAGGGAGTAAAGAGCTCATGAACAATGGCA	2592
QY	1261	AGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAACAACAG	1320
Db	2593	AGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAACAACAG	2652
QY	1321	CCAAAAATCCTGAGATCCCTGGAAGTTCGAGTGTGAGTCTCTGTTACAAAGAGCTTT	1380
Db	2653	CCAAAAATCCTGAGATCCCTGGAAGTTCGAGTGTGAGTCTCTGTTACAAAGAGCTTT	2712
QY	1381	GGATAACATGAACTTCAAAGTGGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAGTGCCCA	1440
Db	2713	GGATAACATGAACTTCAAAGTGGAGTGAACTTCGGAAGAAAGTCTCTCAACATTAGTGCCCA	2772
QY	1441	TTTGGGAAGCAGTTCGACAGTGGGAAGGTCGACCTTCTCTGCGAGGAATCTTGST	1500
Db	2773	TTTGGGAAGCAGTTCGACAGTGGGAAGGTCGACCTTCTCTGCGAGGAATCTTGST	2832
QY	1501	GTGGCTACAGCTGAAGATCATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTCC	1560
Db	2833	GTGGCTACAGCTGAAGATCATGAATTAAGCCGGCAGGCACCTATTGGAGGGCAGCTTCC	2892
QY	1561	AGCAGTTCAAGAGCAAAACGATGACATAGGGGCTTCAAGAGGGAATTGAAACTAAAGA	1620
Db	2893	AGCAGTTCAAGAGCAAAACGATGACATAGGGGCTTCAAGAGGGAATTGAAACTAAAGA	2952
QY	1621	ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTTCACAGAGCAGGCTTTGGA	1680
Db	2953	ACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTTCACAGAGCAGGCTTTGGA	3012
QY	1681	AGGACTAGAGAAACTCTACAGGAGCCACAGAGCTGCCTCTGAGGAGAGAGCCAGAA	1740
Db	3013	AGGACTAGAGAAACTCTACAGGAGCCACAGAGCTGCCTCTGAGGAGAGAGCCAGAA	3072
QY	1741	TGTCACCTCGGCTTACGAAAGCAGCTGAGGAGGTCAATACCTGAGTGGGAAAATTGAA	1800
Db	3073	TGTCACCTCGGCTTACGAAAGCAGCTGAGGAGGTCAATACCTGAGTGGGAAAATTGAA	3132
QY	1801	CCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGAATCCAGGAAT	1860
Db	3133	CCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGAATCCAGGAAT	3192
QY	1861	TCAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGCCAGAGCTGAGGTGATCAAGGGATC	1920
Db	3193	TCAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGCCAGAGCTGAGGTGATCAAGGGATC	3252
QY	1921	CTGGCAGCCGTGGGGCATCTTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA	1980
Db	3253	CTGGCAGCCGTGGGGCATCTTCCTCATTTGACTCTCTCCAAGATCACCTCGAGAAAAGTCAA	3312
QY	1981	GGCACTTCGAGGAGAAATTGGCCTCTGAAAGAGAACGTGAGGCCACGTCAATGACCTTGC	2040
Db	3313	GGCACTTCGAGGAGAAATTGGCCTCTGAAAGAGAACGTGAGGCCACGTCAATGACCTTGC	3372
QY	2041	TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGGTATAAAGCTCAGCACTCTGGAAGA	2100
Db	3373	TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGGTATAAAGCTCAGCACTCTGGAAGA	3432
QY	2101	C 2101	
Db	3433	C 3433	

RESULT 8
AAD37263
ID AAD37263 standard; DNA; 4848 BP.

RESULT 8
AAD37263
ID AAD

XX		AD37263;	
AC			
XX			
XX		21-AUG-2002 (first entry)	
DT			
XX			
XX		Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.	
DE			
XX			
KW		Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;	
KW		adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;	
KW		Becker muscular dystrophy; db.	
XX			
OS		Chimeric - Homo sapiens.	
OS		Chimeric - Cytomegalovirus.	
OS		Chimeric - Unidentified.	
XX			
FN		WO200183595-A2.	
XX			
PD		08-NOV-2001.	
XX			
PF		27-APR-2001; 2001WO-US13677.	
XX			
PR		28-APR-2000; 2000US-20077P.	
XX		(XIAO)/ XIAO X.	
EA			
XX		Xiao X;	
PI			
XX		WPI; 2002-049342/06.	
DR			
XX		New dystrophin minigene for treating Duchenne or Becker muscular	
PT		dystrophy comprises an N-terminal domain or modified N-terminal domain,	
PT		rod repeats, H1 and H4 domains and a cysteine rich domain of a	
PT		dystrophin gene -	
PT			
XX		Example 1; Page 68-70; 71pp; English.	
PS			
XX		The present invention relates to an isolated nucleotide sequence encoding	
CC		a dystrophin minigene. The minigene comprises N-terminal or modified	
CC		N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4	
CC		domains and cysteine-rich domains of dystrophin or utrophin genes. The	
CC		invention also relates to a recombinant adeno-associated virus (AAV)	
CC		comprising dystrophin minigene operably linked to an expression control	
CC		element. The dystrophin minigene in operable linkage with an expression	
CC		control element, in a recombinant adeno-associated virus or retrovirus is	
CC		useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular	
CC		dystrophy (BMD) in a mammalian subject. The present sequence is AAV	
CC		vector plasmid construct containing human dystrophin minigenes, a	
CC		cytomegalovirus (CMV) promoter and a small polyA signal sequence.	
XX			
SQ		Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;	
	Query Match	68.7%; Score 1443; DB 24; Length 4848;	
	Best Local Similarity	84.6%; Pred. No. 0;	
	Matches 1777; Conservative	0; Mismatches	0; Indels 324; Gaps 1;
QY	1	GAGCTATGCCTACACACAGGCGTGTTATGTCAACCACCTCTGACCCCTACACGGAGCCCCATT	60
Db	1580	GAGTATGCCCTACACACAGGCGTGTTATGTCAACCACCTCTGACCCCTACACGGAGCCCCATT	1739
QY	61	TCCITTCAGCATTTGGAAGCCTCTGAAGACAAGTCATTGGCGAGTTTCATTGTGGAGAG	120
Db	1740	TCCITTCAGCATTTGGAAGCCTCTGAAGACAAGTCATTGGCGAGTTTCATTGTGGAGAG	1799
QY	121	TGAAGTAACCTCGACCGTTATCAACACAGCTTTAGAAGAGTATTCGTGGCTTCCTTC	180
Db	1800	TGAAGTAACCTCGACCGTTATCAACACAGCTTTAGAAGAGTATTCGTGGCTTCCTTC	1859
QY	181	TGCTGAGACACATTTGAAGCACAAAGGAGAGATTTCCTAATGATGTGAAGTGGTGAAGA	240
Db	1860	TGCTGAGACACATTTGAAGCACAAAGGAGAGATTTCCTAATGATGTGAAGTGGTGAAGA	1919
QY	241	CCAGTTTCATACTATGAGGGGTACATGTGGATTTTCACACCCCATCAGGGCCGGGTGG	300

Db 1920 CCAGTTTCATCTCATGAGGGCTACATGATGANTTGACAGCCCATCAGGCGCGGTGG 1979
QY 301 TAAATCTTACAAATGGGAGTAAAGCTGATGGACAGCAAAATTTATCAGAAGATGAAGA 360
Db 1980 TAAATCTTACAAATGGGAGTAAAGCTGATGGACAGCAAAATTTATCAGAAGATGAAGA 2039
QY 361 AACTGAAGTACAAAGACAGATGAATCTCTCTAAATCAAGATGGGAATCCCTCAGGTAGC 420
Db 2040 AACTGAAGTACAAAGACAGATGAATCTCTCTAAATCAAGATGGGAATCCCTCAGGTAGC 2099
QY 421 TAGCTGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAACT 480
Db 2100 TAGCTGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAACT 2159
QY 481 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAAACAAAGAAATGGAGGAAGA 540
Db 2160 GAAAGAGTTGAATGACTGGCTAACAAAACAGAGAAAGAAACAAAGAAATGGAGGAAGA 2219
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTACAAACATTAAGTGTCTTCA 600
Db 2220 GCCTCTTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTACAAACATTAAGTGTCTTCA 2279
QY 601 AGAAGATCTAGAACAAAGAACTAGGCTCAATCTCTCACTCAATGCTGCTGCTAGT 660
Db 2280 AGAAGATCTAGAACAAAGAACTAGGCTCAATCTCTCACTCAATGCTGCTGCTAGT 2339
QY 661 TGATGAATCTAGTGGAGATCACGCACTGCTGCTTGGAGAAACAACTTAAGTATTGGG 720
Db 2340 TGATGAATCTAGTGGAGATCACGCACTGCTGCTTGGAGAAACAACTTAAGTATTGGG 2399
QY 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCCGCTGCTTCTTTTACAAGACAT 780
Db 2400 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCCGCTGCTTCTTTTACAAGAC-- 2457
QY 781 CTTTCTCAATGGCAACGCTTACTCAAGAACAGTGCCTTTTATGTCATGCTTTCAGA 840
Db 2458 ----- 2457
QY 841 AAAAAAGATGCAAGTACAAAGATTCACACACTGCTTTAAAGATCAAAATGAATGTT 900
Db 2458 ----- 2457
QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCAT 960
Db 2458 ----- 2457
QY 961 GGGCAAACTGTATTCATCTCAAAACAAAGATCTTCTTCAACTCAAGAAATAGTCACTGAC 1020
Db 2458 ----- 2457
QY 1021 CCAGAAGACGAGATGCTGGATGATCTTGGCGGTGTTGGGATAATTTAGTCCAAAA 1080
Db 2458 ----- 2457
QY 1081 ACTTGAAGAGTACAGACAGATCTAGATTAAGTACCAAGTTCCTCCCTGACCTGGA 1140
Db 2458 ----- ACTCATAGATTACTGCAACAGTTCCTCCCTGACCTGGA 2495
QY 1141 AAAGTTCTTCCCTGGCTTACAGAGCTGAAACAACTGCAATGCTCTACAGGATGCTAC 1200
Db 2496 AAAGTTCTTCCCTGGCTTACAGAGCTGAAACAACTGCAATGCTCTACAGGATGCTAC 2555
QY 1201 CCCTGAAGAAAGGCTCTTGAAGACTCAAGGGAGTAAAGAGCTGATGAAGAAAGTGGCA 1260
Db 2556 CCCTGAAGAAAGGCTCTTGAAGACTCAAGGGAGTAAAGAGCTGATGAAGAAAGTGGCA 2615
QY 1261 AGACTTCAAGGTGAATTTGAAGCTCACAGATGTTTATCAACAACTGATGAAGAAACAG 1320
Db 2616 AGACTTCAAGGTGAATTTGAAGCTCACAGATGTTTATCAACAACTGATGAAGAAACAG 2675
QY 1321 CCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGATGATGATGATGATGATGATGAT 1380
Db 2676 CCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGATGATGATGATGATGATGATGAT 2735

QY 1381 GGATAACATGAATCTCAAGTGAAGTGAATCTCGGAAAAAGTCTCTCAACATTAGTCCCA 1440
Db 2736 GGATAACATGAATCTCAAGTGAAGTGAATCTCGGAAAAAGTCTCTCAACATTAGTCCCA 2795
QY 1441 TTTTGGAGCCAGTCTGACCAAGTGAAGGCTTGACACCTTTCTCTGACAGAACTTCTGCT 1500
Db 2796 TTTTGGAGCCAGTCTGACCAAGTGAAGGCTTGACACCTTTCTCTGACAGAACTTCTGCT 2855
QY 1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGCACTTATTTGGAGGGGACTTTC 1560
Db 2856 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGAGGCACTTATTTGGAGGGGACTTTC 2915
QY 1561 AGCAGTTTCAAGACAGACGATGATAGAGGCTTCAAGAGGAAATTAAGAACTAAAGA 1620
Db 2916 AGCAGTTTCAAGACAGACGATGATAGAGGCTTCAAGAGGAAATTAAGAACTAAAGA 2975
QY 1621 ACCTGTAATCATGAGTACTCTTGAGACTCTGACGAATTTTCTGACAGAGAGGCTTTTGA 1680
Db 2976 ACCTGTAATCATGAGTACTCTTGAGACTCTGACGAATTTTCTGACAGAGAGGCTTTTGA 3035
QY 1681 AGACTAGAGAAACTCTACAGAGGCGCAGAGAGTGCCTCCTTGAGAGAGAGAGCCAGAA 1740
Db 3036 AGACTAGAGAAACTCTACAGAGGCGCAGAGAGTGCCTCCTTGAGAGAGAGAGCCAGAA 3095
QY 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGA 1800
Db 3096 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTTGA 3155
QY 1801 CTTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAGACTCCAGGAACT 1860
Db 3156 CTTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAGACTCCAGGAACT 3215
QY 1861 TCAAGAGGCGCAGGATGAGTGCACCTCAAGCTGCGCCAGCTGAGTGCATCAAGGGATC 1920
Db 3216 TCAAGAGGCGCAGGATGAGTGCACCTCAAGCTGCGCCAGCTGAGTGCATCAAGGGATC 3275
QY 1921 CTGCAAGCCGCTGGGCACTCTCTCATTTGATCTCTTCCAAAGTACCTCGAGAAAGTCAA 1980
Db 3276 CTGCAAGCCGCTGGGCACTCTCTCATTTGATCTCTTCCAAAGTACCTCGAGAAAGTCAA 3335
QY 1981 GGCACCTTCGAGGAGAAATTCGCTCTGAAAGAGAACGTCGAGCACTCAATGACCTTGC 2040
Db 3336 GGCACCTTCGAGGAGAAATTCGCTCTGAAAGAGAACGTCGAGCACTCAATGACCTTGC 3395
QY 2041 TCGCCAGCTTACCACTTTGGGCACTTACGCTCTCAAGTATACCTGAGCACTCTGAGAGA 2100
Db 3396 TCGCCAGCTTACCACTTTGGGCACTTACGCTCTCAAGTATACCTGAGCACTCTGAGAGA 3455
QY 2101 C 2101
Db 3456 C 3456

RESULT 9

AAD37264 standard; DNA; 5060 BP.

XX AAD37264;

XX AC

XX XX

XX 21-AUG-2002 (first entry)

XX Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

XX Chimeric - Homo sapiens.

XX Chimeric - Cytomegalovirus.

XX Chimeric - Unidentified.

XX PN

XX WC00183695-A2.

08-NOV-2001.
27-APR-2001; 2001WO-US13677.
28-APR-2000; 2000US-200777P.
(XIAO/) XIAO X.
Xiao X;
WPI; 2002-049342/06.
New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -
Example 1; Page 70-71; 71pp; English.
The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter and a small polyA signal sequence.
Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;
Query Match 68.7%; Score 1443; DB 24; Length 5060;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 177; Conservative 0; Mismatches 0; Indels 324; Gaps 1;
1 GAGTATGCTACACAGAGCTGTTATGTCACACCTCTGACCCCTACAGGAGCCATT 60
1892 GAGTATGCTACACAGAGCTGTTATGTCACACCTCTGACCCCTACAGGAGCCATT 1951
61 TCCTTACACAGCTTGAAGCTCTGAGACAGTCAATTTGGCAGTTCAATGATGAGAG 120
1952 TCCTTACACAGCTTGAAGCTCTGAGACAGTCAATTTGGCAGTTCAATGATGAGAG 2011
121 TGAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 180
2012 TGAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 2071
181 TGCTGAGGACACATTGACACACAGAGAGATTTCTAATGATGTGAGATGTGGAAGA 240
2072 TGCTGAGGACACATTGACACACAGAGAGATTTCTAATGATGTGAGATGTGGAAGA 2131
241 CCAGTTTCATCTCATGAGGGTACATGATGGATTTTACAGCCCATCAGGCGGGTGG 300
2132 CCAGTTTCATCTCATGAGGGTACATGATGGATTTTACAGCCCATCAGGCGGGTGG 2191
301 TAATATTCTCAATTTGGAGAGTAACTGATGTTGGAACAGGAAATTTATCAGAGATGAAGA 360
2192 TAATATTCTCAATTTGGAGAGTAACTGATGTTGGAACAGGAAATTTATCAGAGATGAAGA 2251
361 AACTGAGTACACAGCAGATGATCTCTTAAATTCAGATGGATGCTCAGGCTAGC 420
2252 AACTGAGTACACAGCAGATGATCTCTTAAATTCAGATGGATGCTCAGGCTAGC 2311
421 TAGCATGAAAAACAAAGCAATTTACATAGATTTTAAAGTATTCAGAAATCAGAAACT 480
2312 TAGCATGAAAAACAAAGCAATTTACATAGATTTTAAAGTATTCAGAAATCAGAAACT 2371
481 GAAAGAGTTGATGACTGGCTTAAACAAAAACAGAGAGAAAGCAAGGAAATGCGGAAGA 540

Db 2372 GAAAGAGTTGAATGACTGGCTAAACAAAAACAGAGAAAGAACAGGAAATGCGGAGAG 2431
Qy 541 GCCTCTTGACCTGATCTTGAGACACCTAAAAACCCCAAGTACAAACAATAGTGTCTTCA 600
Db 2432 GCCTCTTGACCTGATCTTGAGACACCTAAAAACCCCAAGTACAAACAATAGTGTCTTCA 2491
Qy 601 AGAAGATCTAGAACAGAACCAAGTCAAGGTCAATTTCTCTCACTCAATGTTGGTGTAGT 660
Db 2492 AGAAGATCTAGAACAGAACCAAGTCAAGGTCAATTTCTCTCACTCAATGTTGGTGTAGT 2551
Qy 661 TGATGAATCTAGTGGAGATCACCAACTGCTGCTTTGGAAGAACCACTTAAAGTATTGGG 720
Db 2552 TGATGAATCTAGTGGAGATCACCAACTGCTGCTTTGGAAGAACCACTTAAAGTATTGGG 2611
Qy 721 AGATCGATGGGCAAAACATCTGTGATGAGACAGAACCCGCTGGTCTTTTACAAGCAT 780
Db 2612 AGATCGATGGGCAAAACATCTGTGATGAGACAGAACCCGCTGGTCTTTTACAAGCAT 2669
Qy 781 CCTTCTCAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTGATGCGCTTTCAGA 840
Db 2670 ----- 2669
Qy 841 AAAAGAGATGCAAGTGAACAGATTTCACACACTGGCTTTTAAAGATCAAAATGAATGTT 900
Db 2670 ----- 2669
Qy 901 ATCAAGTCTTCAAAACTGCGCTTTTAAAGCGGATCTAGAAAAGAAAGCAATCCAT 960
Db 2670 ----- 2669
Qy 961 GGGCAAACTGTATTCACTCAACCAAGATCTTTTCAACACTGAAGAAATAGTCAGTGAC 1020
Db 2670 ----- 2669
Qy 1021 CCAGAGAGGAGAACGATGGCTGGATAACTTTGGCCGCTGTTGGGATAATTTAGTCCAAA 1080
Db 2670 ----- 2669
Qy 1081 ACTTGAAGAGTACAGACAGACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 1140
Db 2670 ----- ACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2707
Qy 1141 AAAGTTTCTTGCCTGGTTACAGAACTGAAACAACTGCGCAATGTTCTACAGATGCTAC 1200
Db 2708 AAAGTTTCTTGCCTGGTTACAGAACTGAAACAACTGCGCAATGTTCTACAGATGCTAC 2767
Qy 1201 CCGTAAAGAGAGCTCTCTAGAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAAATGGCA 1260
Db 2768 CCGTAAAGAGAGCTCTCTAGAGACTCCAAAGGGAGTAAAGAGCTGATGAAACAAATGGCA 2827
Qy 1261 AGACCTCCAAAGTGAATTTGAAGCTCACAGATGTTTATCACAACTGGATGAAACAG 1320
Db 2828 AGACCTCCAAAGTGAATTTGAAGCTCACAGATGTTTATCACAACTGGATGAAACAG 2887
Qy 1321 CCAAAAAATCTGAGATCCCTGGAAGCTTCGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 2888 CCAAAAAATCTGAGATCCCTGGAAGCTTCGATGATGATGATGATGATGATGATGATGATGAT 2947
Qy 1381 GGATAACATGAATTTCAAGTGGAGTGAATTTGGAAAAAGTCTCTCAACATTTAGTCCCA 1440
Db 2948 GGATAACATGAATTTCAAGTGGAGTGAATTTGGAAAAAGTCTCTCAACATTTAGTCCCA 3007
Qy 1441 TTTGAAAGCCAGTCTGACAGTGAAGCGTCTGCACCTTTCTCTGAGGAGACTTCTGGT 1500
Db 3008 TTTGAAAGCCAGTCTGACAGTGAAGCGTCTGCACCTTTCTCTGAGGAGACTTCTGGT 3067
Qy 1501 GTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGCAGCTTATTTGAGGCGGACTTTCC 1560
Db 3068 GTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGCAGCTTATTTGAGGCGGACTTTCC 3127
Qy 1561 AGCAGTTTGAAGCAGAGACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAACCTAAGA 1620

Db 3128 AGCAGTTCAAGACGAGACGATGTATACATAGGCGCTTCAAGAGGGAAATTTGAAACATAAGA 3187
QY 1621 ACCTGTAATCATGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 1680
Db 3188 ACCTGTAATCATGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGA 3247
QY 1681 AGGACTTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCGAA 1740
Db 3248 AGGACTTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCGAA 3307
QY 1741 TGTCACTCGCTTCTACGAAGCAGCTGAGAGGTCAATCTAGTGGGAAAATTGAA 1800
Db 3308 TGTCACTCGCTTCTACGAAGCAGCTGAGAGGTCAATCTAGTGGGAAAATTGAA 3367
QY 1801 CTTGCACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGAACT 1860
Db 3368 CTTGCACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGAACT 3427
QY 1861 TCAAGAGGCCAGGATGAGCTGGAACCTCAAGCTGGCCAGCTGAGGTGATCAAGGGATC 1920
Db 3428 TCAAGAGGCCAGGATGAGCTGGAACCTCAAGCTGGCCAGCTGAGGTGATCAAGGGATC 3487
QY 1921 CTGGCAGCCGCTGGCGATCTCTCTATTGACTCTCTCAAGATCACTCTCGAGAAAGTCAA 1980
Db 3488 CTGGCAGCCGCTGGCGATCTCTCTATTGACTCTCTCAAGATCACTCTCGAGAAAGTCAA 3547
QY 1981 GGCACCTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGCTGAGCCAGCTCAATGACCTTGC 2040
Db 3548 GGCACCTCGAGGAGAAATTGGCGCTCTGAAAGAGAACGCTGAGCCAGCTCAATGACCTTGC 3607
QY 2041 TCGCCAGCTTACCACCTTGGGCATTTCAGCTCTCACCGTATACCTCAGACACTCTGGAAGA 2100
Db 3608 TCGCCAGCTTACCACCTTGGGCATTTCAGCTCTCACCGTATACCTCAGACACTCTGGAAGA 3667
QY 2101 C 2101
Db 3668 C 3668
RESULT 10
ABK82000
ID ABK82000 standard; DNA; 8689 BP.
XX AC ABK82000;
XX AC
XX AC
DT 13-AUG-2002 (first entry)
XX DE DNA encoding mini-dystrophin protein deltaH2-R19.
XX KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
XX KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200229056-A2.
XX PD 11-APR-2002.
XX PF 04-OCT-2001; 2001WO-US31126.
XX PR 06-OCT-2000; 2000US-238848P.
XX PA (UNMI) UNIV MICHIGAN.
XX PI Chamberlain JS, Harper SQ;
XX DR WPI; 2002-435334/46.
XX PT A composition for preparing therapeutic drugs, has a mini-dystrophin
XX PT peptide comprising a specific number of spectrin-like repeat domains,
XX PT or a nucleic acid sequence encoding the mini-dystrophin peptide -

PS Disclosure; Fig 15; 145pp; English.

XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.

XX SQ Sequence 8689 BP; 2721 A; 1804 C; 1861 G; 2303 T; 0 other;

Query Match 61.5%; Score 1293; DB 24; Length 8689;

Best Local Similarity 72.5%; Pred. NO. 0;

Matches 2101; Conservative 0; Mismatches 0; Indels 798; Gaps 1;

QY 1 GAGCTATGCTACACACAGGCTGCTTATGTCCACCTCTGACCTTACACGAGCCCAAT 60
Db 1099 GAGCTATGCTACACACAGGCTGCTTATGTCCACCTTACACGAGCCCAAT 1158
QY 61 TCCTTCAACAGATTGGAAGCTCTCTGAAGACAGTCTATTTGGCAGTTTATGATGAGAG 120
Db 1159 TCCTTCAACAGATTGGAAGCTCTCTGAAGACAGTCTATTTGGCAGTTTATGATGAGAG 1218
QY 121 TGAAGTAAACCTGACCGCTTATCAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 180
Db 1219 TGAAGTAAACCTGACCGCTTATCAACAGCTTTAGAAAGATTTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTAATGATGTGGAAGTGTGAAGA 1338
QY 241 CCAGTTTCACTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCCGGTTGG 300
Db 1339 CCAGTTTCACTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCCGGTTGG 1398
QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGAGAAATTTATCAGAAATGAAGA 360
Db 1399 TAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGAGAAATTTATCAGAAATGAAGA 1458
QY 361 AACTGAAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 1518
QY 421 TAGCATGGAATAACAAAGCAATTTACATAGATGATTTAATGATCTCCAGAAATCAGAACT 480
Db 1519 TAGCATGGAATAACAAAGCAATTTACATAGATGATTTAATGATCTCCAGAAATCAGAACT 1578
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAGAGAAATGGAGGAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAGAGAAATGGAGGAAGA 1638
QY 541 GCCTCTTGACCTGATCTTGAGAGCTTAAACGCCCAAGTCAACAAATAGGTTCTTCA 600
Db 1639 GCCTCTTGACCTGATCTTGAGAGCTTAAACGCCCAAGTCAACAAATAGGTTCTTCA 1698
QY 601 AGAAGATCTAGAACAAAGCAATGTCAGGGTCAATTTCTCTCACTCATGTTGGTGTAGT 660
Db 1699 AGAAGATCTAGAACAAAGCAATGTCAGGGTCAATTTCTCTCACTCATGTTGGTGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCAGCACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCAGCACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
Db 1819 AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 CTTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGCTTTCAGA 840

Db 1879 CTTCTCAAAATGGCAACGCTTTACTAGAACAGTGGCTTTTATTAGTGCAATGGCTTTCAGA 1938
Qy 841 AAAAGAAATGCAAGTGAACAAAGATTACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1939 AAAAGAAATGCAAGTGAACAAAGATTACACAACCTGGCTTTTAAAGATCAAAATGAAATGTT 1998
Qy 901 ATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTAGAAAGAAAAGCAATCCAT 960
Db 1999 ATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTAGAAAGAAAAGCAATCCAT 2058
Qy 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTCAAGAAATAAAGTCAAGTGAC 1020
Db 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTTTCAACACTCAAGAAATAAAGTCAAGTGAC 2118
Qy 1021 CCAGAAAGCGAAGCATGGCTGGATATTAATTTGGCCGGTGTGGGATTAATTTAGTCCAAA 1080
Db 2119 CCAGAAAGCGAAGCATGGCTGGATATTAATTTGGCCGGTGTGGGATTAATTTAGTCCAAA 2178
Qy 1081 ACTTGAAGAGTACACACAG 1102
Db 2179 ACTTGAAGAGTACACACAGATTTCCAGCAGCCTGACCTAGCTCCTGGACTGACCAC 2238
Qy 1103 ----- 1102
Db 2239 TATTGGAGCCTCTCCTACTACAGACTGTTACTCTGTTGACACAACCTGTGTTACTAAGGA 2298
Qy 1103 ----- 1102
Db 2299 AACTGCCATCTCCAACTAGAAATGCCATCTTCTGTTGATTTGGAGTACCTGCTCTGGC 2358
Qy 1103 ----- 1102
Db 2359 AGATTTCAACCGGCTTGGACAGAACTTACCGACTGGCTTTCTCTGCTTGATCAAGTTAT 2418
Qy 1103 ----- 1102
Db 2419 AAAATCACAGAGGTGATGGTGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCA 2478
Qy 1103 ----- 1102
Db 2479 GAAGCAACAATGCAAGATTGGAAACAGAGCGCTGCCCACTTGGAGAACTCAATTACCGC 2538
Qy 1103 ----- 1102
Db 2539 TGCCCAAAATTTGAAAACAAGACCAGCAATCAAGAGGCTAGAACAACTATTACGGATCG 2598
Qy 1103 ----- 1102
Db 2599 AATTGAAAGAAATTCAGAAATCAGTGGGATGAAGTACAAGAACACCTTCAGAAACCGGAGCA 2658
Qy 1103 ----- 1102
Db 2659 ACAGTTGAATGAATGTTAAAGGATTCAACACATGGCTGGAAGCTTAAGGAAGAGCTGA 2718
Qy 1103 ----- 1102
Db 2719 GCAGTCTTTAGGACAGGCCAGGCAAGCTTGAATCATGGAAGAGGGTCCCTATACGT 2778
Qy 1103 ----- 1102
Db 2779 AGATGCAATCCAAAGAAATCAAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTG 2838
Qy 1103 ----- 1102
Db 2839 GCAGACAAATGTAGATGTGGCAATGACTTGGCCCTGAAACTTCTCCGGGATTTATTCTGC 2898
Qy 1103 ----- 1102
Db 2899 AGATGATACCAAGAAAGTCCACATGATTAACAGAGAAATATCAATGCCCTCTTGGAGAACAT 2958
Qy 1103 ----- 1122
Db 2959 TCATAAAGGTTGAGTGGAGCGAGGCTGCTTTTGGAGAACTCATAGATTACTTGCAACA 3018

RESULT 11
AAD06794
ID

AAD06794 standard; DNA; 5952 BP.

Qy 1123 GTTCCCTCTGACCTGGAAAAGTTTCTTGCTGCTTACAGAGCTGAAAACACTGGCAA 1182
Db 3019 GTTCCCTCTGACCTGGAAAAGTTTCTTGCTGCTTACAGAGCTGAAAACACTGGCAA 3078
Qy 1183 TGTCTCTCAGGATGCTACCCGTAAAGAAAAGCTCTTAGAAGACTCCCAAGGAGTAAAGA 1242
Db 3079 TGTCTCTCAGGATGCTACCCGTAAAGAAAAGCTCTTAGAAGACTCCCAAGGAGTAAAGA 3138
Qy 1243 GCTGATGAAAACAATGGCGAAGACCTCCAAAGTGAAATTTGAAGCTCACAGATGTTATCA 1302
Db 3139 GCTGATGAAAACAATGGCGAAGACCTCCAAAGTGAAATTTGAAGCTCACAGATGTTATCA 3198
Qy 1303 CAACCTGATCAAAACAGCCCAAAAATCTCAGATCCCTGGAAAGTTCCGATGATGAGT 1362
Db 3199 CAACCTGATCAAAACAGCCCAAAAATCTCAGATCCCTGGAAAGTTCCGATGATGAGT 3258
Qy 1363 CCTGTTACAAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAATTCGGAAAAAGTC 1422
Db 3259 CCTGTTACAAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAATTCGGAAAAAGTC 3318
Qy 1423 TCTCAACATTAGTCCCAATTTGGAAAGCCAGTTCTGACCAAGTGGAAAGCTCTGCACCTTTC 1482
Db 3319 TCTCAACATTAGTCCCAATTTGGAAAGCCAGTTCTGACCAAGTGGAAAGCTCTGCACCTTTC 3378
Qy 1483 TCTGCAAGAACTTTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACC 1542
Db 3379 TCTGCAAGAACTTTGCTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACC 3438
Qy 1543 TATTGGAGGCACTTTCCAGCAGTTTCAAGCAAGTTCAGAACAGACGATGATAGGGCTTCAAGAG 1602
Db 3439 TATTGGAGGCACTTTCCAGCAGTTTCAAGCAAGTTCAGAACAGACGATGATAGGGCTTCAAGAG 3498
Qy 1603 GGAATTTGAAACTTAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCT 1662
Db 3499 GGAATTTGAAACTTAAGAAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCT 3558
Qy 1663 GACAGAGCAGCTTTTGGAAAGACTAGAGAAACTCTAACAAGAGCCGAGAGTGCCTCC 1722
Db 3559 GACAGAGCAGCTTTTGGAAAGACTAGAGAAACTCTAACAAGAGCCGAGAGTGCCTCC 3618
Qy 1723 TGAGGAGAGCCGAGATGTCACCTCGCTTCTACGAAAGCAGGCTGAGGAGTCAATAC 1782
Db 3619 TGAGGAGAGCCGAGATGTCACCTCGCTTCTACGAAAGCAGGCTGAGGAGTCAATAC 3678
Qy 1783 TGAGTGGGAAAATTTGAACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCT 1842
Db 3679 TGAGTGGGAAAATTTGAACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCT 3738
Qy 1843 TGAAGACTCCAGGAACCTTCAAGAGGCCACGATGAGTGCACCTCAAGCTGCGCCAGC 1902
Db 3739 TGAAGACTCCAGGAACCTTCAAGAGGCCACGATGAGTGCACCTCAAGCTGCGCCAGC 3798
Qy 1903 TGAGTGTATCAAGGATCCTGGCAGCCGTTGGCGGATCTCCTCATTTGACTCTCTCCAGA 1962
Db 3799 TGAGTGTATCAAGGATCCTGGCAGCCGTTGGCGGATCTCCTCATTTGACTCTCTCCAGA 3858
Qy 1963 TCACCTCGAGAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTCTGAAAGAGACGTGAG 2022
Db 3859 TCACCTCGAGAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTCTGAAAGAGACGTGAG 3918
Qy 2023 CCACTCAATGACCTTTCGCGGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAA 2082
Db 3919 CCACTCAATGACCTTTCGCGGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAA 3978
Qy 2083 CCTCAGCACTCTGGAAGAC 2101
Db 3979 CCTCAGCACTCTGGAAGAC 3997

XX AAD06794;
AC
XX
DT 06-AUG-2001 (first entry)
XX
DE Human dystrophin gene (Becker form).
XX
KW Human; dystrophin; extein; intein; trans-splicing; gene therapy;
KW Duchenne muscular dystrophy; Becker muscular dystrophy; DMD; BMD; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_recomb 2847..2848
FT /*tag= a
FT /label= S4_junction_site
FT /note= "Dystrophin gene is split at this site and
FT attached to an intein sequence. The resulting
FT fragment is used to produce plasmid PSD4"
FT 2952..2953
FT misc_recomb
FT /*tag= b
FT /label= S3_junction_site
FT /note= "Dystrophin gene is split at this site and
FT attached to an intein sequence. The resulting
FT fragment is used to produce plasmid PSD3"
FT 3198..3199
FT misc_recomb
FT /*tag= c
FT /label= S2_junction_site
FT /note= "Dystrophin gene is split at this site and
FT attached to an intein sequence. The resulting
FT fragment is used to produce plasmid PSD2"
FT 3300..3301
FT misc_recomb
FT /*tag= d
FT /label= S1_junction_site
FT /note= "Dystrophin gene is split at this site and
FT attached to an intein sequence. The resulting
FT fragment is used to produce plasmid PSD1"
XX
FN W0200129243-A1.
XX
PD 26-APR-2001.
XX
XX 13-OCT-2000; 2000MO-CA01216.
XX
XX 15-OCT-1999; 99US-0159868.
XX
XX (UYDA-) UNIV DALHOUSIE.
XX (UYPI-) UNIV PITTSBURGH.
XX
XX Paul XL, Xiao X;
XX
XX WPI; 2001-367297/38.
XX
XX Use of spontaneous or automatic protein splicing to join two or more
XX peptides at junction site involves expressing extein peptides having
XX co-reacting portions of split intein attached to them, so that peptides
XX splice
XX
XX Example 1; Fig 2; 81pp; English.
XX
XX The invention relates to a method directed to the use of spontaneous
XX or automatic protein trans-splicing to join two or more peptides at
XX junction site. This method involves expressing extein peptides having
XX co-reacting portions of split intein attached to them, so that peptides
XX will splice automatically under suitable conditions. The invention also
XX provides methods for circumventing virion packaging size limitations in
XX recombinant virus particle, by splitting a coding region for a protein
XX to be delivered into two or more extein genes, which are packaged in
XX separate virus particles and are co-delivered in a target cell for the
XX expression and for subsequent trans-splicing to form the complete
XX protein. In particular, the method is used for trans-splicing human
XX dystrophin and in gene therapies of recombinant adeno-associated virus
XX (AAV) particles that encode trans-spliced dystrophin, for treating

CC diseases such as Duchenne muscular dystrophy (DMD) or Becker muscular
CC dystrophy (BMD). The present sequence is human dystrophin gene
CC (Becker form).
XX
XX Sequence 5952 BP; 1860 A; 1344 C; 1410 G; 1338 T; 0 other;
SQ

Query Match 53.8%; Score 1131; DB 22; Length 5952;
Best Local Similarity 68.8%; Pred. No. 1.5e-305; Indels 960; Gaps 1;
Matches 2101; Conservative 0; Mismatches 0;

QY 1 GAGCTATGCTCACACAGGCTGCTTATGTCTACACCTCTGACCTCTACCGAGGCCATT 60
DB 891 GAGCTATGCTCACACAGGCTGCTTATGTCTACACCTCTGACCTCTACCGAGGCCATT 950
QY 61 TCCTTCACAGCATTGGAAGCTCTCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAG 120
DB 951 TCCTTCACAGCATTGGAAGCTCTCTGAAGACAAGTCATTTGGCAGTTCATTGATGGAG 1010
QY 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTTCTTC 180
DB 1011 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTTCTTC 1070
QY 181 TGTGAGGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 240
DB 1071 TGTGAGGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 1130
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTCG 300
DB 1131 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTTCG 1190
QY 301 TAATATTCTACAAATGGGAAGTAAAGTGAAGTGAACAGGAAATTTATCAGAGATGAAGA 360
DB 1191 TAATATTCTACAAATGGGAAGTAAAGTGAAGTGAACAGGAAATTTATCAGAGATGAAGA 1250
QY 361 AACTGAAGTACAGAGCAGATGATCTCTTAATTCAGATGGATGCTCAGGGTAGC 420
DB 1251 AACTGAAGTACAGAGCAGATGATCTCTTAATTCAGATGGATGCTCAGGGTAGC 1310
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAAATCAGAAACT 480
DB 1311 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTTAAATGGATCTCCAGAAACT 1370
QY 481 GAAAGATTGAATGACTGGCTTAAACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 1371 GAAAGATTGAATGACTGGCTTAAACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1430
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCCAAGTACACAAATAGGTGCTTCA 600
DB 1431 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCCCAAGTACACAAATAGGTGCTTCA 1490
QY 601 AGAAGATCTAGAACAAAGACAAAGTCAGGGTCAATTTCTCCTCCTCATGTTGGTGTAGT 660
DB 1491 AGAAGATCTAGAACAAAGACAAAGTCAGGGTCAATTTCTCCTCCTCATGTTGGTGTAGT 1550
QY 661 TGATGAATCTAGTGGAGATCAGCCTGCTGCTTTGGAGAGAGAGAGAGAGAGAGAGAG 720
DB 1551 TGATGAATCTAGTGGAGATCAGCCTGCTGCTTTGGAGAGAGAGAGAGAGAGAGAGAG 1610
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTCAAGACAT 780
DB 1611 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTCAAGACAT 1670
QY 781 CCTTCTCAATGGAGAGTCTTACTGAGAGACAGTGCCTTTTGTAGTCATGGCTTTCAGA 840
DB 1671 CCTTCTCAATGGAGAGTCTTACTGAGAGACAGTGCCTTTTGTAGTCATGGCTTTCAGA 1730
QY 841 AAAAGAGATGCTGAGTGAACAAAGATTACACAACTGGCTTTTAAAGATCAAAATGAATGTT 900
DB 1731 AAAAGAGATGCTGAGTGAACAAAGATTACACAACTGGCTTTTAAAGATCAAAATGAATGTT 1790
QY 901 ATCAAGTCTTCAAAACATGGCCGCTTTTAAAGCGGATCTAGAAAGAGAGAGAGAGAGATCCAT 960
DB 1791 ATCAAGTCTTCAAAACATGGCCGCTTTTAAAGCGGATCTAGAAAGAGAGAGAGAGATCCAT 1850

QY 961 GGGCAAACTGTAATTCACCTCAAAACAAGATCTTCTTCAACACCTGAAGATAAGTCAGTGAC 1020
Db 1851 GGGCAAACTGTAATTCACCTCAAAACAAGATCTTCTTCAACACCTGAAGATAAGTCAGTGAC 1910
QY 1021 CAGAGACGGAAGCATGGCTGGATTAACCTTGGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
Db 1911 CCAGAGACGGAAGCATGGCTGGATTAACCTTGGCCCGGTGTGGGATAATTTAGTCCAAAA 1970
QY 1081 ACTTGAAGAGTACAGACAG----- 1102
Db 1971 ACTTGAAGAGTACAGACAGGAACTGAAATAGCAGTTCAGCTAAACAACCGGATGT 2030
QY 1103 ----- 1102
Db 2031 GGAAGAGATTTGTCTAAAGGCGAGCATTTGTACAAAGGAAAAACCAAGCCACTCAGCCAGT 2090
QY 1103 ----- 1102
Db 2091 GAAGAGAAATTAGAGATCTGAGCTCTGAGTGGGAAGGCGGTAAACCGTTTACTTCAAGA 2150
QY 1103 ----- 1102
Db 2151 GCTGAGGCAAGCAGCCTGACCTAGCTCCTGGACTGACCCTATTGGAGCCTCTCCTAC 2210
QY 1103 ----- 1102
Db 2211 TCAGACTGTTACTCTGTGTGACACAAACCTGTGTACTTAAGGAACTGCCATCTCCAACT 2270
QY 1103 ----- 1102
Db 2271 AGAAATGCCATCTTCTGTATGTTGGAGTACCTGCTCTGGCAGATTCAAACCGGCTTG 2330
QY 1103 ----- 1102
Db 2331 GACAGAACTTACGACTGGCTTTCTCTGTTGATCAAGTTATATAATTCACAGAGGTGAT 2390
QY 1103 ----- 1102
Db 2391 GGTGGTGACCTTGAGGATATCAACGAGATGATCATCAGCAGAGGCAACAATGCAGGA 2450
QY 1103 ----- 1102
Db 2451 TTTGGAACAGAGGCGTCCCGAGTTGGAGAACTCATTTACCGCTGCCCAAAATTTGAAAAA 2510
QY 1103 ----- 1102
Db 2511 CAAGACCAGCAATCAAGAGGCTAGAACAAATCATTAAGGATCGAATTGAAGAATTAGAA 2570
QY 1103 ----- 1102
Db 2571 TCAGTGGATGAAGTACAGNACACTTCAGAACCGGAGGCAACAGTTGAATGAATGTT 2630
QY 1103 ----- 1102
Db 2631 AAAGGATTCACACAATGGCTGGAAGCTAAGGAAGAGCTCAGCAGGTCTTAGGACAGGC 2690
QY 1103 ----- 1102
Db 2691 CAGAGCCAAGCTTGATGATGGAAGGAGGGTCCCTATACAGTAGATGCAATCCAAAGAA 2750
QY 1103 ----- 1102
Db 2751 AATCAGAGAAACCAAGCAGTTGGCCCAAGACCTCCGCCAGTGGCAGACAAATGTAGATG 2810
QY 1103 ----- 1102
Db 2811 GGCAAATGACTTGGCCCTGAAACTTCTCCGGATTAATTCGAGATGATACCAGAAAAAGT 2870
QY 1103 ----- 1102
Db 2871 CCACATGATAACAGAGAAATCAATGCTCTTGGAGAGCAATTCATAAAAGGCTGAGTGA 2930

QY 1103 -----ACTCATAGATTACTGCAACAGTGTCCCTCGACCTGGACCTGGA 1140
Db 2931 GCGAGAGGCTGCTTTGGAAGAAACTCATAGATTACTGCAACAGTGTCCCTCGACCTGGA 2990
QY 1141 AAAGTTTCTTCCCTGGCTTACAGAGCTGAAACAACTGCCAAATGTCTTACAGGATGCTAC 1200
Db 2991 AAAGTTTCTTCCCTGGCTTACAGAGCTGAAACAACTGCCAAATGTCTTACAGGATGCTAC 3050
QY 1201 CCCTAAGGAAAGGCTCTCTAGAAGCTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCA 1260
Db 3051 CCCTAAGGAAAGGCTCTCTAGAAGCTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCA 3110
QY 1261 AGACTCCAGAGTGAATTTGAAGCTCACAGATGTTTATCAACAACCTGGATGAAACAG 1320
Db 3111 AGACTCCAGAGTGAATTTGAAGCTCACAGATGTTTATCAACAACCTGGATGAAACAG 3170
QY 1321 CCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTTACAAAGACGTTT 1380
Db 3171 CCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGATGCGAGTCTCTTACAAAGACGTTT 3230
QY 1381 GGATAACATGAATTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 3231 GGATAACATGAATTTCAAGTGGAGTGAACTTCGGAAAAAGTCTCTCAACATTAGGTCCCA 3290
QY 1441 TTTGGAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACTTCTCTGCGAGGAACTTCTGCT 1500
Db 3291 TTTGGAAGCCAGTTCTGACAGTGGAAAGCGTCTGCACTTCTCTGCGAGGAACTTCTGCT 3350
QY 1501 GTGGCTACAGCTGAAAAGATGATGAATTAAGCCGCGAGGCACTTATTGGAGCGCACTTCC 1560
Db 3351 GTGGCTACAGCTGAAAAGATGATGAATTAAGCCGCGAGGCACTTATTGGAGCGCACTTCC 3410
QY 1561 AGCAGTTTCAGAGCAGACAGATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAAGA 1620
Db 3411 AGCAGTTTCAGAGCAGACAGATGTACATAGGGCTTCAAGAGGGAATTTGAAAACTAAAGA 3470
QY 1621 ACCTGTATCATAGTACTCTTTGAGACTGTACGAAATATTTCTGACAGCAGCCTTTGGA 1680
Db 3471 ACCTGTATCATAGTACTCTTTGAGACTGTACGAAATATTTCTGACAGCAGCCTTTGGA 3530
QY 1681 AGGACTAGAGAACTCTTACAGAGGCCAGAGACTGCTCCTCCTGAGGAGAGGCCAGAA 1740
Db 3531 AGGACTAGAGAACTCTTACAGAGGCCAGAGACTGCTCCTCCTGAGGAGAGGCCAGAA 3590
QY 1741 TGTCACTCGGCTTTCTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGAAAAATTGAA 1800
Db 3591 TGTCACTCGGCTTTCTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGAAAAATTGAA 3650
QY 1801 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 1860
Db 3651 CCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAAGACTCCAGGAACT 3710
QY 1861 TCAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGTGAATCAAGGGATC 1920
Db 3711 TCAGAGGCCACGGATGAGCTGGACCTCAAGCTCGCCCAAGCTGAGTGAATCAAGGGATC 3770
QY 1921 CTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCAAGATCACTTCGAGAAAGTCAA 1980
Db 3771 CTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCAAGATCACTTCGAGAAAGTCAA 3830
QY 1981 GGCACCTTCGAGGAGAAATTCGCTCTGAAAGAGAAAGTGGAGCCACCTCATGACCTTGC 2040
Db 3831 GGCACCTTCGAGGAGAAATTCGCTCTGAAAGAGAAAGTGGAGCCACCTCATGACCTTGC 3890
QY 2041 TCGCCAGCTTACCCTTTGGGCATTTGAGCTCTCAACCGTATAACCTCAGCACTCTGGAGA 2100
Db 3891 TCGCCAGCTTACCCTTTGGGCATTTGAGCTCTCAACCGTATAACCTCAGCACTCTGGAGA 3950
QY 2101 C 2101
Db 3951 C 3951

RESULT 12
AAD37260
ID AAD37260 standard; DNA; 4414 BP.
XX
AC AAD37260;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-3447.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN WC200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-20077P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 65-66; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
SQ Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 other;
Query Match 53.0%; Score 1114; DB 24; Length 4414;
Best Local Similarity 97.0%; Pred. No. 7.5e-301;
Matches 1135; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
1 GAGGTGCTTACACAGAGCTCTTATGTCACACCTCTGACCCCTACACGAGGCCATT 60
1657 GAGGTGCTTACACAGAGCTCTTATGTCACACCTCTGACCCCTACACGAGGCCATT 1716
61 TCCTTCACAGCATTTGGAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTCAATGATGAGAG 120
1717 TCCTTCACAGCATTTGGAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTCAATGATGAGAG 1776
121 TGAAGTAACTGGACCGTTTCAACAGCTTTAGAGAGTATTATCGTGGCTCTTTC 180
1777 TGAAGTAACTGGACCGTTTCAACAGCTTTAGAGAGTATTATCGTGGCTCTTTC 1836
181 TGCTGAGACACATTGCAAGCAACAGGAGAGATTCTTAATGATGCTGGAAGTGGTGAAGA 240
1837 TGCTGAGACACATTGCAAGCAACAGGAGAGATTCTTAATGATGCTGGAAGTGGTGAAGA 1896

QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTTGG 300
DB 1897 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTTGG 1956
QY 301 TAATATTTCTACAAATTTGGGAAGTAAGCTGAATGGAAACAGGAAATTTATCAGAAGATGAAGA 360
DB 1957 TAATATTTCTACAAATTTGGGAAGTAAGCTGAATGGAAACAGGAAATTTATCAGAAGATGAAGA 2016
QY 361 AACTGAAGTACAAAGGACAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGTTAGC 420
DB 2017 AACTGAAGTACAAAGGACAGATGAATCTCTTAATTTCAAGATGGGAATGCCTCAGGTTAGC 2076
QY 421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480
DB 2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 2136
QY 481 GAAAGAGTTGAATGATCTGCTAAACAAAACAGAAACAGAAACAGAAATTCAGGAAAGA 540
DB 2137 GAAAGAGTTGAATGATCTGCTAAACAAAACAGAAACAGAAATTCAGGAAAGA 2196
QY 541 GCCTCTTGACCTGATCTTTGAAGACCTTAAAGCCCAAGTACAAACATAAGGTGCTTCA 600
DB 2197 GCCTCTTGACCTGATCTTTGAAGACCTTAAAGCCCAAGTACAAACATAAGGTGCTTCA 2256
QY 601 AGAAGATCTAGAAACAAAGTCAAGGTCAATTTCTCACTCACATGCTGCTGCTAGT 660
DB 2257 AGAAGATCTAGAAACAAAGTCAAGGTCAATTTCTCACTCACATGCTGCTGCTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCAGCGACTGCTCTTTGGGAAGCAACATTAAGGTATTGG 720
DB 2317 TGATGAATCTAGTGGAGATCAGCGACTGCTCTTTGGGAAGCAACATTAAGGTATTGG 2376
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGAAACCCCTGGGTTCTTTTCAAGACAT 780
DB 2377 AGATCGATGGGCAACATCTGTAGATGGACAGAAACCCCTGGGTTCTTTTCAAGACAT 2436
QY 781 CTTTCTCAATGGCAACGTTCTTACTGAAGAACAGTCAGTGCCTTTTGTAGTGCATGGCTTTCAGA 840
DB 2437 CTTTCTCAATGGCAACGTTCTTACTGAAGAACAGTCAGTGCCTTTTGTAGTGCATGGCTTTCAGA 2496
QY 841 AAAAGAGATGACAGTCAACAAAGATTCAACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
DB 2497 AAAAGAGATGACAGTCAACAAAGATTCAACAACTGGCTTTTAAAGATCAAAATGAAATGTT 2556
QY 901 ATCAAGTCTTCAAAACTGGCCCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
DB 2557 ATCAAGTCTTCAAAACTGGCCCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2616
QY 961 GGGCAAACTGTATTCTCAACAAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 1020
DB 2617 GGGCAAACTGTATTCTCAACAAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGAC 2676
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAAATTTAGTCCAAA 1080
DB 2677 CCAGAAGACGGAAGCATGGCTGGATAAATTTGCCGGTGTGGGATAAATTTAGTCCAAA 2736
QY 1081 ACTTGAAGAGATACAGACAGATCTAGATTAATCTGCAACAGTTCCTCCCTGACCTGGA 1140
DB 2737 ACTTGAAGAGATACAGACAGATCTAGATTAATCTGCAACAGTTCCTCCCTGACCTGGA 2796
QY 1141 AAAGTCTTCTTGGCTGGCTTACAGAAGCTGA 1170
DB 2797 TGAGCTGGACCTCAAGCTGGCCAGCTGA 2826
RESULT 13
ABK81997
ID ABK81997 standard; DNA; 5417 BP.
XX
AC ABK81997;
XX
DT 13-AUG-2002 (first entry)
XX

DE DNA encoding mini-dystrophin protein deltaR4-R23.
 XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
 KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX WO200229056-A2.
 XX
 XX 11-APR-2002.
 XX
 XX 04-OCT-2001; 2001WO-US31126.
 XX
 XX 06-OCT-2000; 2000US-238848P.
 XX
 XX (UNMI) UNIV MICHIGAN.
 XX
 XX Chamberlain JS, Harper SQ;
 XX
 XX WPI; 2002-435334/46.
 XX
 XX
 XX A composition for preparing therapeutic drugs, has a mini-dystrophin
 PT peptide comprising a specific number of spectrin-like repeat domains,
 PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
 XX
 XX Disclosure; Fig 12; 145pp; English.
 XX
 XX The invention describes a composition comprising a mini-dystrophin
 CC peptide comprising a spectrin-like repeat domain, where the domain
 CC comprises n spectrin-like repeats, and contains no more than n
 CC spectrin-like repeats, where n is an even number between 4-24, or a
 CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
 CC peptide or the polynucleotide encoding it is useful as a medicament,
 CC for preparing a drug for therapeutic application and in the preparation
 CC of a composition for treatment of muscle disease, e.g. Duchenne's
 CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
 CC sequence of the invention.
 XX
 XX Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 other;
 SQ
 Query Match 52.5%; Score 1103.4; DB 24; Length 5417;
 Best Local Similarity 99.9%; Pred. No. 7.8e-298;
 Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 GAGCTATGCTACACACAGCGCTGTTATGTCACACCTCTGACCCCTACACGAGCCCAT 60
 1099 GAGCTATGCTACACACAGCGCTGTTATGTCACACCTCTGACCCCTACACGAGCCCAT 1158
 61 TCCTTTCAGCATTTGGAAGCTCTGAGACAAAGTCAATTTGGCAGTTCATTTGATGGAGAG 120
 1159 TCCTTTCAGCATTTGGAAGCTCTGAGACAAAGTCAATTTGGCAGTTCATTTGATGGAGAG 1218
 121 TGAAGTAAACCTGGACCGCTTATCAACACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTC 180
 1219 TGAAGTAAACCTGGACCGCTTATCAACACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTC 1278
 181 TGCTGAGGACACATTTGGAAGCTCTGAGACAAAGTCAATTTGATGGAGTGTGGAAGA 240
 1279 TGCTGAGGACACATTTGGAAGCTCTGAGACAAAGTCAATTTGATGGAGTGTGGAAGA 1338
 241 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300
 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1398
 301 TAATATTCTCAATTTGGGAAGTAGTGTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 360
 1399 TAATATTCTCAATTTGGGAAGTAGTGTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 1458
 361 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCTCAGGGTAGC 420
 1459 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTTCAAGATGGGAATGCTCAGGGTAGC 1518

QY 421 TAGCATGGAAGAAACAAAGCAATTTTACATAGAGTTTATGATCTCCAGAACTCAGAACT 480
 DB 1519 TAGCATGGAAGAAACAAAGCAATTTTACATAGAGTTTATGATCTCCAGAACTCAGAACT 1578
 QY 481 GAAAGAGTTGATGACTGGCTTAAACAAACAGAAAGAAAGCAAGGAAATGAGGAAGA 540
 DB 1579 GAAAGAGTTGATGACTGGCTTAAACAAACAGAAAGAAAGCAAGGAAATGAGGAAGA 1638
 QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACAAACATAGGTGCTTCA 600
 DB 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACAAACATAGGTGCTTCA 1698
 QY 601 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATCTCTCACTCAGATGTTGGTGTAGT 660
 DB 1699 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATCTCTCACTCAGATGTTGGTGTAGT 1758
 QY 661 TGATCAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 720
 DB 1759 TGATCAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 1818
 QY 721 AGATCGATGGGCAAAACATCTGTAGATGAGACAGAACCGCTGGTCTCTTTTACAGACAT 780
 DB 1819 AGATCGATGGGCAAAACATCTGTAGATGAGACAGAACCGCTGGTCTCTTTTACAGACAT 1878
 QY 781 CCTTCTCAAAATGGCAACGCTCTTACTGAAAGCAAGTGCCTTTTGTAGTCATCGCTTTTCA 840
 DB 1879 CCTTCTCAAAATGGCAACGCTCTTACTGAAAGCAAGTGCCTTTTGTAGTCATCGCTTTTCA 1938
 QY 841 AAAAGAGATGAGTGAACAAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
 DB 1939 AAAAGAGATGAGTGAACAAAGATTCACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
 QY 901 ATCAAGTCTTCAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAGCAATCCAT 960
 DB 1999 ATCAAGTCTTCAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAGCAATCCAT 2058
 QY 961 GGGCAAACTGTATTCTACTCAAAACAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGAC 1020
 DB 2059 GGGCAAACTGTATTCTACTCAAAACAGATCTTCTTCAACACTGAAGAAATAAGTCAGTGAC 2118
 QY 1021 CCAGAAAGCGGAGACATGGCTGATTAACCTTGGCCGCTGTGGGATAATTTAGTCCAAAA 1080
 DB 2119 CCAGAAAGCGGAGACATGGCTGATTAACCTTGGCCGCTGTGGGATAATTTAGTCCAAAA 2178
 QY 1081 ACTTGAAGAGAGTACAGACAGACT 1105
 DB 2179 ACTTGAAGAGAGTACAGACAGACT 2203
 RESULT 14
 ABK82005
 ID ABK82005 standard; DNA; 11241 BP.
 XX
 AC ABK82005;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE cDNA encoding human dystrophin, full length HDMD.
 XX
 KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
 KW Duchenne's muscular dystrophy; DMD; dystrophin; human; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200229056-A2.
 XX
 XX 11-APR-2002.
 XX
 XX 04-OCT-2001; 2001WO-US31126.
 XX
 XX 06-OCT-2000; 2000US-238848P.
 XX
 XX (UNMI) UNIV MICHIGAN.
 PA

XX Chamberlain JS, Harper SQ;
PI WPI; 2002-435334/46.
XX
XX
XX A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
XX Example 2; Fig 23; 145pp; English.
XX
XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a human dystrophin
CC polynucleotide sequence used in the creation of the mini-dystrophin
CC peptides of the invention.
XX
XX Sequence 11241 BP; 3738 A; 2325 C; 2656 G; 2522 T; 0 other;
Query Match 52.5%; Score 1103.4; DB 24; Length 11241;
Best Local Similarity 99.9%; Pred. No. 1.2e-297;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 GAGCTATGCTACACAGAGCTCTTATCTCACCACCTCTGACCTTACACGAGCCCAT 60
1099 GAGCTATGCTACACAGAGCTCTTATCTCACCACCTCTGACCTTACACGAGCCCAT 1158
61 TCCTTCACAGCATTTGGAAGCTCTTCAAGACAGTCACTTTGGCAGTTCATTGATGAGAG 120
1159 TCCTTCACAGCATTTGGAAGCTCTTCAAGACAGTCACTTTGGCAGTTCATTGATGAGAG 1218
121 TGAAGTAACTCGGACCTTATCAACAGCTTTAGAAAGATTTATCTGCTCTCTTTC 180
1219 TGAAGTAACTCGGACCTTATCAACAGCTTTAGAAAGATTTATCTGCTCTCTTTC 1278
181 TGCTGAGACACATTTGGAAGCTCTTCAAGACAGTCACTTTGGAAGTGGTGAAGA 240
1279 TGCTGAGACACATTTGGAAGCTCTTCAAGACAGTCACTTTGGAAGTGGTGAAGA 1338
241 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 300
1339 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 1398
301 TAATATTTCAATTTGGAAGTAACTGATTTGGAAGTAACTTATCAGAGATGAAGA 360
1399 TAATATTTCAATTTGGAAGTAACTGATTTGGAAGTAACTTATCAGAGATGAAGA 1458
361 AACTGGAAGTAACTGGAAGTAACTGATTTGGAAGTAACTTATCAGAGATGAAGA 420
1459 AACTGGAAGTAACTGGAAGTAACTGATTTGGAAGTAACTTATCAGAGATGAAGA 1518
421 TAGCATGGAAGTAACTGGAAGTAACTGATTTGGAAGTAACTTATCAGAGATGAAGA 480
1519 TAGCATGGAAGTAACTGGAAGTAACTGATTTGGAAGTAACTTATCAGAGATGAAGA 1578
481 GAAAGAGTTGATGATGCTGCTTCAACAAAGTAACTGGAAGTAACTTATCAGAGATGAAGA 540
1579 GAAAGAGTTGATGATGCTGCTTCAACAAAGTAACTGGAAGTAACTTATCAGAGATGAAGA 1638
541 GCCTTTGGACCTGATCTTGAAGTAACTTGAAGTAACTTGAAGTAACTTGAAGTAACTTGA 600
1639 GCCTTTGGACCTGATCTTGAAGTAACTTGAAGTAACTTGAAGTAACTTGAAGTAACTTGA 1698
601 AGAGATCTAGACAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 660
1699 AGAGATCTAGACAGACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1758

QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTCTCTTTGGAGAACAACTTAAGTATTGGG 720
DB 1759 TGATGAATCTAGTGGAGATCAGCAACTGCTCTCTTTGGAGAACAACTTAAGTATTGGG 1818
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGACGCTGGTTCCTTTTACAAGACAT 780
DB 1819 AGATCGATGGGCAACATCTGTAGATGGACAGACGCTGGTTCCTTTTACAAGACAT 1878
QY 781 CCTTCTCAATGCGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGA 840
DB 1879 CCTTCTCAATGCGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTCAGA 1938
QY 841 AAAAGAGATGCGAGTGAACAGATTCACCAACTGCTTTTAAAGTCAAAATGAAATGTT 900
DB 1939 AAAAGAGATGCGAGTGAACAGATTCACCAACTGCTTTTAAAGTCAAAATGAAATGTT 1998
QY 901 ATCAAGTCTTCAAAAAGTGCCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
DB 1999 ATCAAGTCTTCAAAAAGTGCCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCATCTCAAAAGATCTTCTTCAACACTGAAGATAAGTCACTGAC 1020
DB 2059 GGGCAAACTGTATTCATCTCAAAAGATCTTCTTCAACACTGAAGATAAGTCACTGAC 2118
QY 1021 CCAGAGACGGAAGCATGGCTGGATTAACCTTGGCCGCTGTTGGGATAATTTAGTCCAAA 1080
DB 2119 CCAGAGACGGAAGCATGGCTGGATTAACCTTGGCCGCTGTTGGGATAATTTAGTCCAAA 2178
QY 1081 ACTTGAAGAGTACAGCACAGACT 1105
DB 2179 ACTTGAAGAGTACAGCACAGACT 2203
RESULT 15
ABK82002
ID ABR82002 standard; DNA; 11443 BP.
XX
AC ABR82002;
XX
DT 13-AUG-2002 (first entry)
XX
DE DNA encoding mini-dystrophin protein deltaR9-R16.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
XX Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
OS Homo sapiens.
XX Synthetic.
XX
PN WO200229056-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US31126.
XX
PR 06-OCT-2000; 2000US-238848P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Chamberlain JS, Harper SQ;
XX
XX WPI; 2002-435334/46.
XX
XX A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
XX Disclosure; Fig 17; 145pp; English.
XX
XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a

CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX
SQ Sequence 11443 BP; 3707 A; 2339 C; 2502 G; 2895 T; 0 other;
Query Match 52.58; Score 1103.4; DB 24; Length 11443;
Best Local Similarity 99.9%; Pred. No. 1.2e-297;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTCACACCTCTGACCCCTACACGAGCCCAT 60
DB 1099 GAGCTATGCTTACACACAGGCTGCTTATGTCACACCTCTGACCCCTACACGAGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAAGCTCTGAGACAGTCAATTTGGCAGTTCATTTGATGAGAG 120
DB 1159 TCCTTCACAGCATTTGGAAGCTCTGAGACAGTCAATTTGGCAGTTCATTTGATGAGAG 1218
QY 121 TGAAGTAAACCTGACCGCTTATCAACACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 180
DB 1219 TGAAGTAAACCTGACCGCTTATCAACACAGCTTTAGAGAGTATTATCGTGGCTTCTTTC 1278
QY 181 TGTGAGGACACATTTGACACACACAGAGAGATTTCTATGATGTGGAAGTGTGAAAGA 240
DB 1279 TGTGAGGACACATTTGACACACACAGAGAGATTTCTATGATGTGGAAGTGTGAAAGA 1338
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGGCCGGGTGG 300
DB 1339 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTCACAGCCCATCAGGGCCGGGTGG 1398
QY 301 TAATATTCTCAATTTGGGAAGTAAAGCTGATTTGAAGACAGGAAATTTATCAGACATGAA 360
DB 1399 TAATATTCTCAATTTGGGAAGTAAAGCTGATTTGAAGACAGGAAATTTATCAGACATGAA 1458
QY 361 AACTCAAGTACAGAGCAGATGATCTCTAAATCAAGATGGATGCGCTCAGGGTAGC 420
DB 1459 AACTCAAGTACAGAGCAGATGATCTCTAAATCAAGATGGATGCGCTCAGGGTAGC 1518
QY 421 TAGCATGGAACAAACAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAAACT 480
DB 1519 TAGCATGGAACAAACAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAAACT 1578
QY 481 GAAGAGTTGAATGACTGGCTTACAAACAAACAGAGAAAGAACAGAGGAATGAGGAAGA 540
DB 1579 GAAGAGTTGAATGACTGGCTTACAAACAAACAGAGAAAGAACAGAGGAATGAGGAAGA 1638
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTAAACGCGCAAGTACACACATTAAGCTGCTTCA 600
DB 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCGCAAGTACACACATTAAGCTGCTTCA 1698
QY 601 AGAAGATCTAGAACACAGAACAGTCAAGGTCATTTCTCACTCACATGCTGCTGTAGT 660
DB 1699 AGAAGATCTAGAACACAGAACAGTCAAGGTCATTTCTCACTCACATGCTGCTGTAGT 1758
QY 661 TGATGAATCTAGTGGATCAGCAACTGCTGCTTTGGAGAGCAACTTAAGGTATTGGG 720
DB 1759 TGATGAATCTAGTGGATCAGCAACTGCTGCTTTGGAGAGCAACTTAAGGTATTGGG 1818
QY 721 AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACAT 780
DB 1819 AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACAT 1878
QY 781 CTTCTCAAAATGGCAACGCTCTTACTGAGACAGTGGCTTTTGTGATGCTGGCTTTCAGA 840
DB 1879 CTTCTCAAAATGGCAACGCTCTTACTGAGACAGTGGCTTTTGTGATGCTGGCTTTCAGA 1938
QY 841 AAAAGAGATGAGTGAACAGATTTACACAACTGGCTTTTAAAGATCAAAATGAATGTT 900
DB 1939 AAAAGAGATGAGTGAACAGATTTACACAACTGGCTTTTAAAGATCAAAATGAATGTT 1998

Search completed: February 1, 2004, 12:03:57
Job time : 536.045 secs

QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 960
DB 1999 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAAGAAAAAGCAATCCAT 2058
QY 961 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTGAC 1020
DB 2059 GGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACACTGAAGAATAAGTCAAGTGAC 2118
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 1080
DB 2119 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTGGGATAATTTAGTCCAAAA 2178
QY 1081 ACTTGAAGAGTACACACAGCT 1105
DB 2179 ACTTGAAGAGTACACACAGATT 2203

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 1, 2004, 10:14:55 ; Search time 7484.44 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-2_COPY_900_3000
Perfect score: 2101
Sequence: 1 gagctatgcctacacacagg.....acotcagcactctggaagac 2101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
GenEmbl.*			
1:	gb	ba.*	AX538622 Sequence
2:	gb	htg.*	AR304538 Sequence
3:	gb	in.*	AX114289 Sequence
4:	gb	om.*	AX538619 Sequence
5:	gb	ov.*	AX538624 Sequence
6:	gb	pat.*	AX538627 Sequence
7:	gb	ph.*	AX409637 Sequence
8:	gb	pl.*	AX538581 Sequence
9:	gb	pr.*	MI8533 Homo sapien
10:	gb	ro.*	AR220819 Sequence
11:	gb	scs.*	X14298 Human mRNA
12:	gb	sy.*	AX538620 Sequence
13:	gb	un.*	AX538621 Sequence
14:	gb	vi.*	AF070485 Canis fam
15:	em	ba.*	E30220 Shortened d
16:	em	fun.*	MI8025 Mouse dyster
17:	em	hum.*	AX306153 Sequence
18:	em	in.*	AX538582 Sequence
19:	em	mu.*	M68859 Mouse dyster
20:	em	or.*	AR093392 Sequence
21:	em	ov.*	AR142592 Sequence
22:	em	pat.*	E30221 Shortened d
23:	em	ph.*	E30219 Shortened d
24:	em	pl.*	X13369 Chicken rRN
25:	em	ro.*	X06178 Human adult
26:	em	scs.*	X02018 Shortened d
27:	em	sy.*	AX538612 Sequence
28:	em	un.*	AX538583 Sequence
29:	em	vi.*	X69086 H.sapiens m
30:	em	wtg_hum.*	AX538611 Sequence
31:	em	wtg_inv.*	BC024140 Mus muscu
32:	em	wtg_other.*	AF339031 Danio rer
33:	em	wtg_mus.*	BC050183 Mus muscu
34:	em	wtg_pln.*	A63605 Sequence 7
35:	em	wtg_rod.*	AR281528 Sequence 9
36:	em	wtg_mam.*	A63607 Sequence 9
37:	em	wtg_vrt.*	AR281529 Sequence
38:	em	wtg_hum.*	AX107972 Sequence
39:	em	wtg_hum.*	AX538584 Sequence
40:	em	wtg_hum.*	Y12229 M.musculus
41:	em	wtg_hum.*	AX538589 Sequence
42:	em	wtg_hum.*	AX538588 Sequence
43:	em	wtg_hum.*	AX538588 Sequence
44:	em	wtg_hum.*	AX538588 Sequence
45:	em	wtg_hum.*	AX538590 Sequence

ALIGNMENTS

RESULT 1
AX538622
LOCUS AX538622 8689 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

FEATURES		Location/Qualifiers	
source	1..8689	/organism="synthetic construct"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:32630"	
		/note="Synthetic"	
BASE COUNT	2721 a 1804 c 1861 g 2303 t		
ORIGIN			
Query Match 61.5%; Score 1293; DB 6; Length 8689;			
Best local Similarity 72.5%; Pred. No. 3.3e-296;			
Matches 2101; Conservative 0; Mismatches 0; Indels 798; Gaps 1;			
QY	1	GAGCTATGCTACACACAGGCTGCTTATGTCACACCTCTGACCCCTACACGGAGCCCAT	60
DB	1099	GAGCTATGCTACACACAGGCTGCTTATGTCACACCTCTGACCCCTACACGGAGCCCAT	1158
QY	61	TCCTTCACAGCATTTGGAAGCTCTGAACACAGTCATTTGACGATTCATTCATGAGAG	120
DB	1159	TCCTTCACAGCATTTGGAAGCTCTGAACACAGTCATTTGACGATTCATTCATGAGAG	1218
QY	121	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAAGATATTATCGTGGCTTTCTTC	180
DB	1219	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAAGATATTATCGTGGCTTTCTTC	1278
QY	181	TGCTGAGGACACATTTGAGACCAACAGAGAGATTTCTATGATGTGGAAGTGGTGAAGA	240
DB	1279	TGCTGAGGACACATTTGAGACCAACAGAGAGATTTCTATGATGTGGAAGTGGTGAAGA	1338
QY	241	CCAGTTTCATACATCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTGG	300
DB	1339	CCAGTTTCATACATCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGTGG	1398
QY	301	TAATATTTCAATTTGGAAAGTAAAGTGTGGAACAGGAAATTTATCAGAGATGAAGA	360
DB	1399	TAATATTTCAATTTGGAAAGTAAAGTGTGGAACAGGAAATTTATCAGAGATGAAGA	1458
QY	361	AACGTGATACAGAGCAGATGATCTCTTAATTTCAAGTGGATGCTCAGGGTAGC	420
DB	1459	AACGTGATACAGAGCAGATGATCTCTTAATTTCAAGTGGATGCTCAGGGTAGC	1518
QY	421	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCGAATCAGAACT	480
DB	1519	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCGAATCAGAACT	1578
QY	481	GAAAGATTGAATGATCTGGCTTAACAAACAGAGAAAGAAAGAAATGGAGGA	540
DB	1579	GAAAGATTGAATGATCTGGCTTAACAAACAGAGAAAGAAAGAAATGGAGGA	1638
QY	541	GCCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAAGGTGCTTCA	600
DB	1639	GCCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAAGGTGCTTCA	1698
QY	601	AGAGATCTAGAACAGAAACAGTCAGGGTCAATTTCTCCTCCTCAGATGGTGGTAGT	660
DB	1699	AGAGATCTAGAACAGAAACAGTCAGGGTCAATTTCTCCTCCTCAGATGGTGGTAGT	1758
QY	661	TGATGATCTAGTGGAGATCACGCACTGCTGCTTTTGGAAAGAACAACTTAAAGTATTGG	720
DB	1759	TGATGATCTAGTGGAGATCACGCACTGCTGCTTTTGGAAAGAACAACTTAAAGTATTGG	1818
QY	721	AGATCGATGGCAACATCTGTAGATCGACAGAGACCGGTGGGTCTTTTACAGACAT	780
DB	1819	AGATCGATGGCAACATCTGTAGATCGACAGAGACCGGTGGGTCTTTTACAGACAT	1878
QY	781	CTTCTCAATGGCAAGCTTCTTCTAGAGACAGTGGCTTTTGTAGTGCATGGCTTTTCA	840
DB	1879	CTTCTCAATGGCAAGCTTCTTCTAGAGACAGTGGCTTTTGTAGTGCATGGCTTTTCA	1938
QY	841	AAAAGAGATGCAGTGAACAGAGTTTCAACAATCGGCTTTTAAAGATCAAAATGAATGT	900
DB	1939	AAAAGAGATGCAGTGAACAGAGTTTCAACAATCGGCTTTTAAAGATCAAAATGAATGT	1998
QY	901	ATCAAGTCTTCAAAAACTGSCCGCTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT	960
DB	1999	ATCAAGTCTTCAAAAACTGSCCGCTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCAT	2058
QY	961	GGCAGAACTGATTTCACTCAAAACACAGATCTTTTCAACACTGAAGAAATAAGTCACTGAC	1020
DB	2059	GGCAGAACTGATTTCACTCAAAACACAGATCTTTTCAACACTGAAGAAATAAGTCACTGAC	2118
QY	1021	CCAGAAGACGGAAGCATGGCTGGATAACTTTTGGCCCGTGTGGGATAATTTAGTCCAAAA	1080
DB	2119	CCAGAAGACGGAAGCATGGCTGGATAACTTTTGGCCCGTGTGGGATAATTTAGTCCAAAA	2178
QY	1081	ACTTGAAGAGTACAGCACAG	1102
DB	2179	ACTTGAAGAGTACAGCACAGATTTTCAACAGACCTGACCTAGCTCTGGACTGACCAC	2238
QY	1103	-----	1102
DB	2239	TATTGGAGCCTCTCTACTCTCAGACTGTGTTACTCTGTGTGACACACACTGTGTTACTAAGGA	2298
QY	1103	-----	1102
DB	2299	AACCTGCATCTCAAACTAGAAATGCGATCTTCTCTGTGATGTTGGAGGTACCTCTCTGGC	2358
QY	1103	-----	1102
DB	2359	AGATTTCAACCGGGCTTGGACAGAACTTACCAGACTGGCTTCTCTGTGATCAAGTTAT	2418
QY	1103	-----	1102
DB	2419	AAATCACAGAGGTGATGTTGGTGACCTTGGAGGATATCAACAGATGATCATCAAGCA	2478
QY	1103	-----	1102
DB	2479	GAAGCAACAAATGCAGGATTTGGACAGAGCGCTCCCGACTTGGAAAGAACTCATTTACCGC	2538
QY	1103	-----	1102
DB	2539	TGCCAAAATTTGAAAAACCAAGACCAATCAAGAGCTTGAACCAATCATTTACGGATCG	2598
QY	1103	-----	1102
DB	2599	AATTGAAGAAATTCAGAAATCAGTGGGATGAATACAGAACACCTTCAGAACCGGAGCA	2658
QY	1103	-----	1102
DB	2659	ACAGTTGAATGAATGTTAAAGGATTCACACAATGGCTGGAAGCTAAGGAAGAAGCTGA	2718
QY	1103	-----	1102
DB	2719	GCAGTCTTAGGACAGGCCAGCCAGCTTGAATGAGTCAATGGAAGGGTCCCTATACAGT	2778
QY	1103	-----	1102
DB	2779	AGATGCAATCCAAAAGAAATTCACAGAAACCAAGCAGTTGGCCCAAGACCTCCGCCAGTG	2838
QY	1103	-----	1102
DB	2839	GCAGACAAATGTAGATGTGGCAATGACTTGGCCCTGAAACTTCTCCGGGATTTATTTCTGC	2898
QY	1103	-----	1102
DB	2899	AGATGATACAGAAAAGTCCACATGATAACAGAGAAATCAATGCTCTTTGGAGAAGCAT	2958
QY	1103	-----	1122
DB	2959	TCATAAAGGGTGAAGTGAAGGAGGCTTTTGGAAAGAACTCATAGATTACTGCAACA	3018
QY	1123	GTTCCTCCCTGACCTGGAAAGTTTCTTCTGCTGTACAGAGCTGAAACAACTGCGCAA	1182
DB	3019	GTTCCTCCCTGACCTGGAAAGTTTCTTCTGCTGTACAGAGCTGAAACAACTGCGCAA	3078
QY	1183	TGTCCTACAGGATGCTACCCGTAAGGAAAGGCTCTTAGAAGACTCCCAAGGGAGTAAAGA	1242

Db 3079 TGTCTACAGATCTACCGTAGGAAGGCTCTTAGAAGCTCCAGGAGTAAAGA 3138
Qy 1243 GCTGATGAACAAATGGAAGACCTCCAGGTGAAATTTGAAGCTCACACAGATGTTATCA 1302
Db 3139 GCTGATGAACAAATGGAAGACCTCCAGGTGAAATTTGAAGCTCACACAGATGTTATCA 3198
Qy 1303 CAACCTGGATGAACAAAGCAGCAAAATTCCTGAGATCCCTGGAGGTTCCGATGATCAGT 1362
Db 3199 CAACCTGGATGAACAAAGCAGCAAAATTCCTGAGATCCCTGGAGGTTCCGATGATCAGT 3258
Qy 1363 CCTGTTACAAGAGCTTTGGATACATCAACTTCAAGTGGAGTGAATTCGGAAGATGTC 1422
Db 3259 CCTGTTACAAGAGCTTTGGATACATCAACTTCAAGTGGAGTGAATTCGGAAGATGTC 3318
Qy 1423 TCTCAACATTTAGTCCCTTTGGAGGAGCTTTGACACAGTGAAGGCTGTCACCTTTC 1482
Db 3319 TCTCAACATTTAGTCCCTTTGGAGGAGCTTTGACACAGTGAAGGCTGTCACCTTTC 3378
Qy 1483 TCTCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC 1542
Db 3379 TCTCAGGAACCTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACC 3438
Qy 1543 TATTGGAGGAGCTTTCCAGCAGTTTCAGAGCAGAAAGATGATGATAGGAGCTTTCAAGAG 1602
Db 3439 TATTGGAGGAGCTTTCCAGCAGTTTCAGAGCAGAAAGATGATGATAGGAGCTTTCAAGAG 3498
Qy 1603 GGAATTGAAACTAAAGAACTGTAATCATGAGTACTTGGAGACTGTACGAATATTTCT 1662
Db 3499 GGAATTGAAACTAAAGAACTGTAATCATGAGTACTTGGAGACTGTACGAATATTTCT 3558
Qy 1663 GACAGAGCAGCTTTGGAGAGCTTAGAGAACTCTACAGAGCCCAAGAGAGTGCCTCC 1722
Db 3559 GACAGAGCAGCTTTGGAGAGCTTAGAGAACTCTACAGAGCCCAAGAGAGTGCCTCC 3618
Qy 1723 TGAGGAGAGCCAGATGCTCACTCGCTTCTACGAAGCAGGCTCAGAGAGTCAATAC 1782
Db 3619 TGAGGAGAGCCAGATGCTCACTCGCTTCTACGAAGCAGGCTCAGAGAGTCAATAC 3678
Qy 1783 TGAGTGGGAAAAATTTGAACCTGCACTCGCTGCTGAGCAGAGAAAAATAGATGAGACCTT 1842
Db 3679 TGAGTGGGAAAAATTTGAACCTGCACTCGCTGCTGAGCAGAGAAAAATAGATGAGACCTT 3738
Qy 1843 TGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGAGACCTCAAGCTGGCCCAAGC 1902
Db 3739 TGAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGAGACCTCAAGCTGGCCCAAGC 3798
Qy 1903 TGAGTGTATCAAGGATCCTGGCAGCCCGTGGGCGATCTCTCAATTGACTCTCTCCCAAGA 1962
Db 3799 TGAGTGTATCAAGGATCCTGGCAGCCCGTGGGCGATCTCTCAATTGACTCTCTCCCAAGA 3858
Qy 1963 TCACTTCGAGAAAGTCAAGGCACTTTCGAGAGGAAATTTGGCCCTCTGAAAGAGAACGTGAG 2022
Db 3859 TCACTTCGAGAAAGTCAAGGCACTTTCGAGAGGAAATTTGGCCCTCTGAAAGAGAACGTGAG 3918
Qy 2023 CCACGTCATGACCTTGTCTGCGCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAA 2082
Db 3919 CCACGTCATGACCTTGTCTGCGCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATAA 3978
Qy 2083 CCTCAGCACTCTGGAAGAC 2101
Db 3979 CCTCAGCACTCTGGAAGAC 3997

RESULT 2
AR304538
LOCUS AR304538
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM

Unclassified.
REFERENCE 1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
1..5952
/organism="unknown"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 53.8%; Score 1131; DB 6; Length 5952;
Best Local Similarity 68.6%; Pred. No. 1e-257;
Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;
Qy 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCTCACACCTCTGACCTCACCGAGGCCATT 60
Db 891 GAGCTATGCCCTACACACAGGCTGCTTATGTCTCACACCTCTGACCTCACCGAGGCCATT 950
Qy 61 TCCTTTCACAGCATTTTGAAGCTCCTGGAAGACAAGTCAATTCGCAAGTTCATTGATGGAGAG 120
Db 951 TCCTTTCACAGCATTTTGAAGCTCCTGGAAGACAAGTCAATTCGCAAGTTCATTGATGGAGAG 1010
Qy 121 TGAAGTAAACCTGACCGCTTATCAACAGCTTTAGAGAAGTATATCTGCTGCTTTTC 180
Db 1011 TGAAGTAAACCTGACCGCTTATCAACAGCTTTAGAGAAGTATATCTGCTGCTTTTC 1070
Qy 181 TCCTCAGGACACATTTGAAGCACAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
Db 1071 TCCTCAGGACACATTTGAAGCACAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1130
Qy 241 CAGTTTCATCTATGAGGGGTACATGATGATGATTTGACAGCCCATCAGGCCCGGTTGG 300
Db 1131 CAGTTTCATCTATGAGGGGTACATGATGATGATTTGACAGCCCATCAGGCCCGGTTGG 1190
Qy 301 TAATATTCTACAAATTTGGAAGTAAAGCTGATTTGGAACAGAGAAAAATTTATCAGAAGATGAAGA 360
Db 1191 TAATATTCTAATTTGGAAGTAAAGCTGATTTGGAACAGAGAAAAATTTATCAGAAGATGAAGA 1250
Qy 361 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGAGTGGGAATGCCTCAGGGTAGC 420
Db 1251 AACTGAAGTACAGAGCAGATGAATCTCTCTAAATTCAGAGTGGGAATGCCTCAGGGTAGC 1310
Qy 421 TAGCATGGAACCAACAACTTTTACATAGAGTATTTAATGATCTCCAGAAATCAGAAACT 480
Db 1311 TAGCATGGAACCAACAACTTTTACATAGAGTATTTAATGATCTCCAGAAATCAGAAACT 1370
Qy 481 GAAAGAGTTGAATGATCTGCTTAACAAAAACAGAGAAAGAAACAGAGAAAAATGGAGGAAGA 540
Db 1371 GAAAGAGTTGAATGATCTGCTTAACAAAAACAGAGAAAGAAACAGAGAAAAATGGAGGAAGA 1430
Qy 541 GCCTCTTGACCTGATCTTGAAGACCTAAAGCCCAAGTACAAACAATAAGTCTTTCA 600
Db 1431 GCCTCTTGACCTGATCTTGAAGACCTAAAGCCCAAGTACAAACAATAAGTCTTTCA 1490
Qy 601 AGAAGATCTAGAACAAAGAACAGGTGAGGTCAATTTCTCTCACTCACTGGTGGTGGTAGT 660
Db 1491 AGAAGATCTAGAACAAAGAACAGGTGAGGTCAATTTCTCTCACTCACTGGTGGTGGTAGT 1550
Qy 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAGAACACTTAAGTTATGGG 720
Db 1551 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAGAACACTTAAGTTATGGG 1610
Qy 721 AGATCGATGGCAAAACATCTGTAGATGGAAGAGACCGCTGGTCTTTTACAAAGACAT 780
Db 1611 AGATCGATGGCAAAACATCTGTAGATGGAAGAGACCGCTGGTCTTTTACAAAGACAT 1670
Qy 781 CCTTCTCAAATGGCAAGCTTTTACAGAGAACAGTGCCTTTTGTAGTGCATGCTTTTCA 840
Db 1671 CCTTCTCAAATGGCAAGCTTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGCTTTTCA 1730
Qy 841 AAAAGAGATGTCAGTGAACAAAGATTCACACAACTGGCTTTTAAAGATCAAATGAAATGTT 900

1731 AAAAGAGATGTCAGTGAACAGATTCACACACTGGCTTTAAGATCAAAATGAATGTT 1790
901 ATCAAGTCTTCAAAAACGTGGCCGTTTAAAGCGGATCTAGAAAAGAAAGCAATCCAT 960
1791 ATCAAGTCTTCAAAAACGTGGCCGTTTAAAGCGGATCTAGAAAAGAAAGCAATCCAT 1850
961 GGGCAAACTGTATTACACTCAAAACAGATCTTCTTCAACACTGAAGAAATAGTCAGTGAC 1020
1851 GGGCAAACTGTATTACACTCAAAACAGATCTTCTTCAACACTGAAGAAATAGTCAGTGAC 1910
1021 CCAGAAGCGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAAATTTAGTCCAAAA 1080
1911 CCAGAAGCGGAAGCATGGCTGGATAAATTTGCCCGGTGTGGGATAAATTTAGTCCAAAA 1970
1081 ACTTGAAGAGTACAGACACAG----- 1102
1971 ACTTGAAGAGTACAGACACAGAACTGAAATAGCATTCAAGCTAAACACACCGGATGT 2030
1103 ----- 1102
2031 GGAAGAGATTTGTCTAAAGGCGCAGCATTTGTCTCAAGGAAACACAGCCACTCAGCCAGT 2090
1103 ----- 1102
2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGGAAAGCGGTAAACCGTTTACTTCAAGA 2150
1103 ----- 1102
2151 GCTGAGGCAAGCAGCCTGACCTAGCTCCTGGACTGACCACTATTGGAGCCTCTCTAC 2210
1103 ----- 1102
2211 TCAGACTGTTACTCTGGTGACACAACTGTGGTTACTAAGGAAACTGCCACTCTCCAAACT 2270
1103 ----- 1102
2271 AGAATGCCATCTTCTTGATGTTGGAGTACCTGCTCTGGCAGATTTCACCGGGCTTG 2330
1103 ----- 1102
2331 GACAGAACTTACCGATGCGCTTCTGCTTGATCAAGTTATATAATCACAGAGGGTAT 2390
1103 ----- 1102
2391 GGTGGTGACTTGAGGATATCAACGAGATGATCATCAAGCAGAAAGCAACATGCCAGGA 2450
1103 ----- 1102
2451 TTTGGAAACAGAGCGCTCCCGAGTTGGAAGAACTCATTTACCGCTGCCCAAATTTGAAAAA 2510
1103 ----- 1102
2511 CAAGACGCAATCAAGAGGCTAGAACATATTACGGATCGAATTGAAGAAATTCAGAA 2570
1103 ----- 1102
2571 TCAGTGGGATGAAGTACAAGAACACCTTCAGAACCGGCGCAACAGTTGATGAATGTT 2630
1103 ----- 1102
2631 AAAGGATTCAACAAATGGCTGGAAGCTTAAGGAAGAAGCTGAGAGGTTTAGGACAGGC 2690
1103 ----- 1102
2691 CAGAGCCAGCTTGATCATGGAAGAGGGTCCCTATACAGTAGATGCAATCCAAAAGNA 2750
1103 ----- 1102
2751 AATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAAATGTAGATGT 2810
1103 ----- 1102

2811 GCGAAATGACTTGGCCCTGAAACTTCTCCGGATATTCTCGAGATGATACCAAGAAAGT 2870
1103 ----- 1102
2871 CCACATGATACAGAGAATATCAATGCCCTTTGGAGAGCATTCATAAAAGGTGAGTGA 2930
1103 ----- 1102
2931 GCGAGAGGCTGCTTTTGGAAAGAACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGA 2990
1141 AAAGTTTCTTGGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 1200
2991 AAAGTTTCTTGGCTGGCTTACAGAAAGCTGAAACAACTGCCAATGTCTTACAGGATGCTAC 3050
1201 CCGTAAGGAAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAACAAATGGGA 1260
3051 CCGTAAGGAAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAACAAATGGGA 3110
1261 AGACCTCCNAGGTGAATTTGAAGCTCACACAGATGTTTTATCACACCTGGATGAAACAG 1320
3111 AGACCTCCNAGGTGAATTTGAAGCTCACACAGATGTTTTATCACACCTGGATGAAACAG 3170
1321 CCAAAAAATCCTTAGATCCCTGGAAAGTTCGATGATGATGATGATGATGATGATGATGATGAT 1380
3171 CCAAAAAATCCTTAGATCCCTGGAAAGTTCGATGATGATGATGATGATGATGATGATGATGAT 3230
1381 GGATTAACATGAATTCAGTGGAGTGAATTTGGAAAGCTTCGAAAGAAAGTCTCTCAACATTAGTCCCA 1440
3231 GGATTAACATGAATTCAGTGGAGTGAATTTGGAAAGCTTCGAAAGAAAGTCTCTCAACATTAGTCCCA 3290
1441 TTTGGAAGCCAGTTCTGACCAAGTGAAGCTTCGACCTTTCTCTGCAAGAACTTTCTGAT 1500
3291 TTTGGAAGCCAGTTCTGACCAAGTGAAGCTTCGACCTTTCTCTGCAAGAACTTTCTGAT 3350
1501 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
3351 GTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3410
1561 AGCAGTTCAAGAGCAGACGATGTATAGGAGGCTTCAAGAGGAAATGAAACAACTAAAGA 1620
3411 AGCAGTTCAAGAGCAGACGATGTATAGGAGGCTTCAAGAGGAAATGAAACAACTAAAGA 3470
1621 ACCTGTAATCATGAGTACTTTGAGACTGTAGAAATTTCTGACAGAGCAGCCTTTGGA 1680
3471 ACCTGTAATCATGAGTACTTTGAGACTGTAGAAATTTCTGACAGAGCAGCCTTTGGA 3530
1681 AGACTAGAGAACTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
3531 AGACTAGAGAACTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3590
1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGAAATTTGAA 1800
3591 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGAAATTTGAA 3650
1801 CTTGCACTCGGCTGAGTGGCAGAGAAATAGATGAGACCTTGAAGACTCCAGAACT 1860
3651 CTTGCACTCGGCTGAGTGGCAGAGAAATAGATGAGACCTTGAAGACTCCAGAACT 3710
1861 TCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGGTCATCAAGGGATC 1920
3711 TCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGGTCATCAAGGGATC 3770
1921 CTGGCAGCCGCTGGGAGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
3771 CTGGCAGCCGCTGGGAGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3830
1981 GGCACCTTCGAGGAGAAATTCGGCTCTGAAGAGAGACGTCGAGCCACGTCATGACCTTGC 2040
3831 GGCACCTTCGAGGAGAAATTCGGCTCTGAAGAGAGACGTCGAGCCACGTCATGACCTTGC 3890
2041 TCGCCAGCTTACCACTTTGGGCAATTCAGTCTTCAACGATTAACCTCAGCAGCTCTGGAAGA 2100
3891 TCGCCAGCTTACCACTTTGGGCAATTCAGTCTTCAACGATTAACCTCAGCAGCTCTGGAAGA 3950

QY 2101 C 2101
Db 3951 C 3951

RESULT 3
LOCUS AX114289 5952 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent WO0129243.
ACCESSION AX114289
VERSION AX114289.1 GI:14031259
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Paul, X.L. and Xiao, X.
TITLE Method and vector for producing and transferring trans -spliced peptides
JOURNAL Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION (US)
FEATURES
source
1. 5952
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
misc_feature 2897..2898
/note="S4 junction site"
misc_feature 3198..3199
/note="S2 junction site"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN

Query Match 53.8%; Score 1131; DB 6; Length 5952;
Best Local Similarity 68.6%; Pred. No. le-257;
Matches 2101; Conservative 0; Mismatches 0; Indels 960; Gaps 1;

QY 1 GAGCTATGCCCTACACACAGGCTGTTATGTACACACCTCTGACCCCTACACGGAGCCCAIT 60
Db 891 GAGCTATGCCCTACACACAGGCTGTTATGTACACACCTCTGACCCCTACACGGAGCCCAIT 950

QY 61 TCCCTTCACAGATTTGGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTCAITGATGAGAG 120
Db 951 TCCCTTCACAGATTTGGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTCAITGATGAGAG 1010

QY 121 TGAAGTAACTCGACCGCTTATCAACAGCTTTAGAAAGATATATCGTGGCTTCTTTC 180
Db 1011 TGAAGTAACTCGACCGCTTATCAACAGCTTTAGAAAGATATATCGTGGCTTCTTTC 1070

QY 181 TGCTGAGGACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA 240
Db 1071 TGCTGAGGACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAAGA 1130

QY 241 CCAGTTTCATCTCATGAGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1131 CCAGTTTCATCTCATGAGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1190

QY 301 TAATATTCTCAATTTGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1191 TAATATTCTCAATTTGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 1250

QY 361 AACTGAAGTACAAGACGATGATTCCTTAATTTCAAGATGGGAATGCCCTCAGGTAGC 420
Db 1251 AACTGAAGTACAAGACGATGATTCCTTAATTTCAAGATGGGAATGCCCTCAGGTAGC 1310

QY 421 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT 480
Db 1311 TAGCATGAAAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT 1370

QY 481 GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAACAAAGGAAATGGGGAAGA 540

Db 1371 GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAACAAAGGAAATGGGGAAGA 1430
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACACAAACATAGGTGCTTCA 600
Db 1431 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACACAAACATAGGTGCTTCA 1490
QY 601 AGAAGATCTAAGAACAAAGCAAGTCAGGTCATTTCTCACTCACATGGTGGTGTAGT 660
Db 1491 AGAAGATCTAAGAACAAAGCAAGTCAGGTCATTTCTCACTCACATGGTGGTGTAGT 1550
QY 661 TGATGAATCTAGTGGAGATCAACGCACTGCTTGTGGAAGAACAACTTAAGTATTGGG 720
Db 1551 TGATGAATCTAGTGGAGATCAACGCACTGCTTGTGGAAGAACAACTTAAGTATTGGG 1610
QY 721 AGATCGATGGCAACACATCTGTAGATGACAGAGACCGCTGGGTCTTTTACAGACAT 780
Db 1611 AGATCGATGGCAACACATCTGTAGATGACAGAGACCGCTGGGTCTTTTACAGACAT 1670
QY 781 CCTTCTCAATGGCAACGTCCTTACTGAAGAACAGTCGCTTTTAAAGTGCATGGCTTTCAGA 840
Db 1671 CCTTCTCAATGGCAACGTCCTTACTGAAGAACAGTCGCTTTTAAAGTGCATGGCTTTCAGA 1730
QY 841 AAAAGAGATCGATGACAGATTTCAACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
Db 1731 AAAAGAGATCGATGACAGATTTCAACAACTGGCTTTAAAGATCAAAATGAAATGTT 1790
QY 901 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1791 ATCAAGTCTTCAAAAACTGGCCGTTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1850
QY 961 GGGCAACTGTTACTCAACAAAGATCTTCTTCAACACTGAGATAGTACAGTACAGTAC 1020
Db 1851 GGGCAACTGTTACTCAACAAAGATCTTCTTCAACACTGAGATAGTACAGTACAGTAC 1910
QY 1021 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGATTAATTTAGTCCAAA 1080
Db 1911 CCAGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGATTAATTTAGTCCAAA 1970
QY 1081 ACTTGAAAAGAGTACAGCAGCAG----- 1102
Db 1971 ACTTGAAAAGAGTACAGCAGCAGCAGGAAACTGAAATAGCAGTTCAAGCTAAACACCGGATGT 2030
QY 1103 ----- 1102
Db 2031 GGAAGAGATTTTGTCTAAAGCGGAGCATTTGTACAGGAAACCCAGCCACTCAGCCAGT 2090
QY 1103 ----- 1102
Db 2091 GAAGAGGAAGTTAGAAGATCTGAGCTCTGAGTGGAGCGGTAAACCGTTTACTTCAAGA 2150
QY 1103 ----- 1102
Db 2151 GCTGAGGCAAGCAGCCTGACCTAGTCTCTGGACTGACCACTATTGGAGCCTCTCCTAC 2210
QY 1103 ----- 1102
Db 2211 TCAGACTGTTACTCTGGTGCACAACTCTGTTACTAAGGAAACTGCCACTCTCCAAACT 2270
QY 1103 ----- 1102
Db 2271 AGAATGCCATCTTCTTGTATGTTGGAGGTACTGCTCTGGCAGATTTCAACCGGGCTTG 2330
QY 1103 ----- 1102
Db 2331 GACAGACTTACCGACTGCTTCTCTCTGTTGATCAAGTTATATAAATCACAGAGGGTAT 2390
QY 1103 ----- 1102
Db 2391 GGTGGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGAGGCAACAATGCAGGA 2450
QY 1103 ----- 1102

Db 2451 TTTGGAACAGAGCGCTCCCGAGTTGGAAGAACTCATTTACCGCTGCCCAAAATTTGAAAA 2510
Qy 1103 ----- 1102
Db 2511 CAAGACCAGCAATCAAGAGGCTAGAACATCATTTACGGATCGAATTGAAAGATTGAGAA 2570
Qy 1103 ----- 1102
Db 2571 TCAGTGGGATGAAGTACAAAGAACACCTTCAGAACCGGAGGCAACAGTTGAATGAATGTT 2630
Qy 1103 ----- 1102
Db 2631 AAAGGATTCAACACAAATGGCTGGAAGCTAAGGAAGAAGCTGACAGGTTCTTAGACAGGC 2690
Qy 1103 ----- 1102
Db 2691 CAGAGCCAAAGCTTGAGTCATGGAAGAGGCTCCCTATACAGTAGATGCAATCCAAAGAA 2750
Qy 1103 ----- 1102
Db 2751 AATCAGAGAAACCAAGAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAAATGTATGAT 2810
Qy 1103 ----- 1102
Db 2811 GGCAAAATGACTTGGCCCTGAAACTTCTCCGGAGATTATCTGCAGATGATACAGAAAAAGT 2870
Qy 1103 ----- 1102
Db 2871 CCACATGATAACAGAGATATCAATGCCCTTTGGAGAAGCAATCATAAAAGGTCAGTGA 2930
Qy 1103 -----ACTCATAGATTACTGCAACAGATTCCCTCCCTCGACTCGA 1140
Db 2931 GCGAGAGGCTGTTTGGAGAAACTCATAGATTACTGCAACAGATTCCCTCCCTCGACTCGA 2990
Qy 1141 AAAGTTCTTCTGCTGCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGAGTGTAC 1200
Db 2991 AAAGTTCTTCTGCTGCTTACAGAACTGAAACAACTGCCAATGTCTCTACAGAGTGTAC 3050
Qy 1201 CCGTAAGGAAGGCTCTAGAGATCCCAAGAGGAGTAAAGAGCTGATGAACAATGGA 1260
Db 3051 CCGTAAGGAAGGCTCTAGAGATCCCAAGAGGAGTAAAGAGCTGATGAACAATGGA 3110
Qy 1261 AGACCTCCAGAGTGAATTCAGCTACACAGATGTTTATCAACAACCTGGATGAAACAG 1320
Db 3111 AGACCTCCAGAGTGAATTCAGCTACACAGATGTTTATCAACAACCTGGATGAAACAG 3170
Qy 1321 CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAAGTCTCTGTTACAAAGAGCTTT 1380
Db 3171 CCAAAAAATCCTGAGATCCCTGGAAGTTCCGATGATGCAAGTCTCTGTTACAAAGAGCTTT 3230
Qy 1381 GGATAACATGAATTCAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAGGTCCCA 1440
Db 3231 GGATAACATGAATTCAGTGGAGTGAACCTCGGAAAAGTCTCTCAACATTAGGTCCCA 3290
Qy 1441 TTTGGAAGCCAGTTCTGACAGGTGAAGCTCTGCACTTTCTCTGAGAACTTCTGTT 1500
Db 3291 TTTGGAAGCCAGTTCTGACAGGTGAAGCTCTGCACTTTCTCTGAGAACTTCTGTT 3350
Qy 1501 GTGCTCAGCTGAAGATGATGATTAAGCCGAGGACCTATTGGAGCGACCTTCC 1560
Db 3351 GTGCTCAGCTGAAGATGATGATTAAGCCGAGGACCTATTGGAGCGACCTTCC 3410
Qy 1561 AGCAGTTTCAAGAGCAGAACATGATACATAGGCTTCAAGAGGGAATTTGAAAACTAAAGA 1620
Db 3411 AGCAGTTTCAAGAGCAGAACATGATACATAGGCTTCAAGAGGGAATTTGAAAACTAAAGA 3470
Qy 1621 ACCTGTAAATAGTACTCTTTGAGATGTGACGATATTTCTGACAGAGCAGCTTTGGA 1680
Db 3471 ACCTGTAAATAGTACTCTTTGAGATGTGACGATATTTCTGACAGAGCAGCTTTGGA 3530
Qy 1681 AGGACTAGAGAACTCTACAGAGGACCCAGAGAGCTCCCTCCTGAGGAGAGCCAGAA 1740
Db 3531 AGGACTAGAGAACTCTACAGAGGACCCAGAGAGCTCCCTCCTGAGGAGAGAGCCAGAA 3590

Qy 1741 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATATCTAGTGGAAAAAATTGAA 1800
Db 3591 TGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCATATCTAGTGGAAAAAATTGAA 3650
Qy 1801 CTTGCACTCGGCTGAGTGGCAGAGAAAAATAGATGAGACCTTGAAGAACTCCAGGAAT 1860
Db 3651 CTTGCACTCGGCTGAGTGGCAGAGAAAAATAGATGAGACCTTGAAGAACTCCAGGAAT 3710
Qy 1861 TCAAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 1920
Db 3711 TCAAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATCAAGGGATC 3770
Qy 1921 CTGGCAGCCGCTGGGCGATCTCTCTATTGACTCTTCCAAAGATCACTCGAGAACTCAA 1980
Db 3771 CTGGCAGCCGCTGGGCGATCTCTCTATTGACTCTTCCAAAGATCACTCGAGAACTCAA 3830
Qy 1981 GGCACCTTCGAGGAGAAATTCGGCTCTTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGC 2040
Db 3831 GGCACCTTCGAGGAGAAATTCGGCTCTTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGC 3890
Qy 2041 TCGCCAGCTTACCACTTTGGCAATTCAGCTCTCCCGTATACCTCAGCACTCTGGAAGA 2100
Db 3891 TCGCCAGCTTACCACTTTGGCAATTCAGCTCTCCCGTATACCTCAGCACTCTGGAAGA 3950
Qy 2101 C 2101
Db 3951 C 3951

RESULT 4
AX538619
LOCUS AX538619
DEFINITION Sequence 39 from Patent WO0229056.
ACCESSION AX538619
VERSION AX538619.1 GI:25271163
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
Patent: WO 0229056-A 39 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source
1..5417
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"

BASE COUNT 1700 a 1192 c 1182 g 1343 t
ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 5417;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGCTATGCTTACACACAGGCTGCTTATGTCACCCACCTCTGACCCCTACCGAGCCCAAT 60
Db 1099 GAGCTATGCTTACACACAGGCTGCTTATGTCACCCACCTCTGACCCCTACCGAGCCCAAT 1158
Qy 61 TCCCTCAGCATTTTGGAGCTCTGAGACAGTCAATTTGGCAGTTTCAATGATGGAGAG 120
Db 1159 TCCCTCAGCATTTTGGAGCTCTGAGACAGTCAATTTGGCAGTTTCAATGATGGAGAG 1218
Qy 121 TGAAGTAACTCGACCGTTATCAACAGCTTTAGAAAGAGTATTTCGTGGCTTCTTTC 180
Db 1219 TGAAGTAACTCGACCGTTATCAACAGCTTTAGAAAGAGTATTTCGTGGCTTCTTTC 1278
Qy 181 TGCTGAGGACACATTCGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGTGGAAGA 240
Db 1279 TGCTGAGGACACATTCGCAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGTGGAAGA 1398

QY	241	CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCGGGTTGG	300		
DB	1339	CCAGTTTCTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCGGGTTGG	1398		
QY	301	TAAATTTCTCAATTTGGGAAGTAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA	360		
DB	1399	TAAATTTCTCAATTTGGGAAGTAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA	1458		
QY	361	AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCGAATGGGAATGCTCAGGCTAGC	420		
DB	1459	AACTGAAGTACAAGAGCAGATGAATCTCTTAAATTCGAATGGGAATGCTCAGGCTAGC	1518		
QY	421	TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAACT	480		
DB	1519	TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAACT	1578		
QY	481	GAAAGAGTTGAATGACTGGCTTAAACAAAAACAGAGAAAGAAACAAAGGAAATGAGGAAGA	540		
DB	1579	GAAAGAGTTGAATGACTGGCTTAAACAAAAACAGAGAAAGAAACAAAGGAAATGAGGAAGA	1638		
QY	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATTAAGTGGTTCA	600		
DB	1639	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATTAAGTGGTTCA	1698		
QY	601	AGAAGATCTAGAAACAAGAACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	660		
DB	1699	AGAAGATCTAGAAACAAGAACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	1758		
QY	661	TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG	720		
DB	1759	TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGGAAGAACAACTTAAGGTATTGGG	1818		
QY	721	AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTCTCTTTTACAGACAT	780		
DB	1819	AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTCTCTTTTACAGACAT	1878		
QY	781	CCTTCTCAAAATGGCAACCGTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTT	840		
DB	1879	CCTTCTCAAAATGGCAACCGTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTT	1938		
QY	841	AAAAGAAGATGCAAGTGAACAAAGATTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT	900		
DB	1939	AAAAGAAGATGCAAGTGAACAAAGATTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT	1998		
QY	901	ATCAAGTCTTCAAAAACATGGCGCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	960		
DB	1999	ATCAAGTCTTCAAAAACATGGCGCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	2058		
QY	961	GGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACCTGAAGAAATTAAGTCAGTGAC	1020		
DB	2059	GGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACCTGAAGAAATTAAGTCAGTGAC	2118		
QY	1021	CCAGAAAGCGAAGATGGCTGGATTAATTTGCCCGTGTGGGATTAATTTAGTCCAAAA	1080		
DB	2119	CCAGAAAGCGAAGATGGCTGGATTAATTTGCCCGTGTGGGATTAATTTAGTCCAAAA	2178		
QY	1081	ACTTGAAGAGTACAGACAGACT	1105		
DB	2179	ACTTGAAGAGTACAGACAGACT	2203		
RESULT 5	AX538624	11443 bp	DNA	linear	PAT 23-NOV-2002
LOCUS	AX538624				
DEFINITION	Sequence 44 from Patent WO0229056.				
ACCESSION	AX538624				
VERSION	AX538624.1	GI:25271175			
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1				

AUTHORS	Chamberlain,J.S. and Harper,S.Q.									
TITLE	Mini-dystrophin nucleic acid and peptide sequences									
JOURNAL	Patent: WO 0229056-A 44 11-APR-2002;									
	THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)									
FEATURES	Location/Qualifiers									
source	1..11443									
	/organism="synthetic construct"									
	/mol_type="genomic DNA"									
	/db_xref="taxon:32630"									
	/note="Synthetic"									
BASE COUNT	3707 a 2339 c 2502 g 2895 t									
ORIGIN										
	Query Match	52.5%;	Score 1103.4;	DB 6;	Length 11443;					
	Best Local Similarity	99.9%;	Pred. No. 3.6e-251;							
	Matches 1104;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;	
QY	1	GAGCTATGCCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACACGAGGCCCAATT	60							
DB	1099	GAGCTATGCCCTACACACAGGCTGCTTATGTCTACACACCTCTGACCCCTACACGAGGCCCAATT	1158							
QY	61	TCCTTCACAGCATTTGGAAGCTCCTTGAAGCAAGTCAATTTGGCAGTTCAATGATGGAGAG	120							
DB	1159	TCCTTCACAGCATTTGGAAGCTCCTTGAAGCAAGTCAATTTGGCAGTTCAATGATGGAGAG	1218							
QY	121	TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAAGATATATTCGTGGCTTCTTTC	180							
DB	1219	TGAAGTAAACCTGGACCGGTTATCAAAACAGCTTTAGAAGATATATTCGTGGCTTCTTTC	1278							
QY	181	TGCTCAGGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	240							
DB	1279	TGCTCAGGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1338							
QY	241	CCAGTTTTCATCTCATGAGGGTACATGATGGAATTGACAGCCCATCAGGCGCGGTTGG	300							
DB	1339	CCAGTTTTCATCTCATGAGGGTACATGATGGAATTGACAGCCCATCAGGCGCGGTTGG	1398							
QY	301	TAATATTTCTCAATTTGGAAGTAAAGTCAATTTGGAACAGGAAATTTATCAGAAGATCAAGA	360							
DB	1399	TAATATTTCTCAATTTGGAAGTAAAGTCAATTTGGAACAGGAAATTTATCAGAAGATCAAGA	1458							
QY	361	AACTGAAGTCAAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC	420							
DB	1459	AACTGAAGTCAAGAGCAGATGAATCTCTCTAAATTTCAAGATGGGAATGCCCTCAGGGTAGC	1518							
QY	421	TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAAACT	480							
DB	1519	TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAAACT	1578							
QY	481	GAAAGAGTTTGAATGATGCTGCTTAAACAAACAGAGAAAGAAACAAAGGAAATGGAGGAAGA	540							
DB	1579	GAAAGAGTTTGAATGATGCTGCTTAAACAAACAGAGAAAGAAACAAAGGAAATGGAGGAAGA	1638							
QY	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATTAAGTGGTCTTCA	600							
DB	1639	GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATTAAGTGGTCTTCA	1698							
QY	601	AGAAGATCTAGAAACAAAGCAAGTCAAGGTCGAATTCCTCACTCAATGGTGGTGGTAGT	660							
DB	1699	AGAAGATCTAGAAACAAAGCAAGTCAAGGTCGAATTCCTCACTCAATGGTGGTGGTAGT	1758							
QY	661	TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTGGAGAAACAACTTAAGGTATTGGG	720							
DB	1759	TGATGAATCTAGTGGAGATCAGGCAACTGCTGCTTGGAGAAACAACTTAAGGTATTGGG	1818							
QY	721	AGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAAGACAT	780							
DB	1819	AGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAAGACAT	1878							
QY	781	CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCA	840							
DB	1879	CCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTTTTGTAGTCATGGCTTTTCA	1938							

841 AAAAGAGATGAGTGAACAAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
1939 AAAAGAGATGAGTGAACAAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 1998
901 ATCAAGTCTTCAAAACTGGCGGTTTAAAGCGGATCTAGAAAGAAAGCAATCCAT 960
1999 ATCAAGTCTTCAAAACTGGCGGTTTAAAGCGGATCTAGAAAGAAAGCAATCCAT 2058
961 GGGCAAACTGTATTCACTCAAAACAGATCTCTTTCAACACCTGAAGATAAAGTCAGTGAC 1020
2059 GGGCAAACTGTATTCACTCAAAACAGATCTCTTTCAACACCTGAAGATAAAGTCAGTGAC 2118
1021 CCAGAAAGCGGAGCATGGCTGGATAAATCTTCCCGGTTTGGGATATTTAGTCCAAAA 1080
2119 CCAGAAAGCGGAGCATGGCTGGATAAATCTTCCCGGTTTGGGATATTTAGTCCAAAA 2178
1081 ACTTGAAGAGATACAGACAGACT 1105
2179 ACTTGAAGAGATACAGACAGATT 2203

RESULT 6
AX538627
LOCUS AX538627 12057 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 47 from Patent WO229056.
ACCESSION AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.O.
TITLE Mini-dyetrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1..12057
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="Synthetic"

BASE COUNT 4020 a 2456 c 2790 g 2791 t

ORIGIN
Query Match 52.5%; Score 1103.4; DB 6; Length 12057;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GAGCTATGCTACACACAGGCTGCTTATGTCTACCACTCTGACCCCTACACGAGCCCAT 60
1099 GAGCTATGCTACACACAGGCTGCTTATGTCTACCACTCTGACCCCTACACGAGCCCAT 1158
61 TCCTTACAGCATTTGAGAGCTCCGGAAGACAGTCAATTTGGCAGTTCAATTCAGAGAG 120
1159 TCCTTACAGCATTTGAGAGCTCCGGAAGACAGTCAATTTGGCAGTTCAATTCAGAGAG 1218
121 TGAAGTAACTCTGACCGTTTCAACAGCTTTAGAAAGATGATTAATCGTGGCTCTTTC 180
1219 TGAAGTAACTCTGACCGTTTCAACAGCTTTAGAAAGATGATTAATCGTGGCTCTTTC 1278
181 TGCTGAGGACACATTCGAGCAACAGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 240
1279 TGCTGAGGACACATTCGAGCAACAGAGAGATTTCTAATGATGTGGAAAGTGGTGAAGA 1338
241 CCAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGCGCGGGTTGG 300
1339 CCAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGCGCGGGTTGG 1398
301 TAATATTCTACAATTTGGGAAGTAAAGCTGATTTGGAACAGGAAAAATTCAGAGATGAAGA 360
1399 TAATATTCTACAATTTGGGAAGTAAAGCTGATTTGGAACAGGAAAAATTCAGAGATGAAGA 1458

361 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCCTCAGGTAGC 420
1459 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCCTCAGGTAGC 1518
421 TAGCATGGAAAAACAAGCAATTTACATAGAGTCTTTAATGGATCTCCAGAATCAGAAACT 480
1519 TAGCATGGAAAAACAAGCAATTTACATAGAGTCTTTAATGGATCTCCAGAATCAGAAACT 1578
481 GAAAGAGTTGAATGATGGCTTAACAAAAACAAGAAAGCAAGAAAGAAAAATGGAGGAAGA 540
1579 GAAAGAGTTGAATGATGGCTTAACAAAAACAAGAAAGCAAGAAAGAAAAATGGAGGAAGA 1638
541 GCCTCTTGGACCTGATCTTTGAAGACCTTAAGCCCAAGTACAAACAACATAAGGTCTTCA 600
1639 GCCTCTTGGACCTGATCTTTGAAGACCTTAAGCCCAAGTACAAACAACATAAGGTCTTCA 1698
601 AGAAGATCTAGAACAGAAACAGTCAAGGTCAATCTCTCACTCACTGCTGGTGGTGTAGT 660
1699 AGAAGATCTAGAACAGAAACAGTCAAGGTCAATCTCTCACTCACTGCTGGTGGTGTAGT 1758
661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTTAAGGTATTGGG 720
1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTTAAGGTATTGGG 1818
721 AGATCCATGGGCAAAACATCTGTAGATGGAAGAACCGCTCGGTTCTTTTACAAGACAT 780
1819 AGATCCATGGGCAAAACATCTGTAGATGGAAGAACCGCTCGGTTCTTTTACAAGACAT 1878
781 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCA 840
1879 CCTTCTCAAAATGGCAACGCTTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTCA 1938
841 AAAAGAGATGCAAGTGAACAAAGATTACACAACTGGCTTTAAAGATCAAAATGAAATGTT 900
1939 AAAAGAGATGCAAGTGAACAAAGATTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
901 ATCAAGTCTTCAAAACTGGCGGTTTAAAGCGGATCTAGAAAGAAAGAAAGCAATCCAT 960
1999 ATCAAGTCTTCAAAACTGGCGGTTTAAAGCGGATCTAGAAAGAAAGAAAGCAATCCAT 2058
961 GGGCAAACTGTATTCACTCAACAAAGATCTCTTTCAACACTGGAAGAAATAAGTCAGTGAC 1020
2059 GGGCAAACTGTATTCACTCAACAAAGATCTCTTTCAACACTGGAAGAAATAAGTCAGTGAC 2118
1021 CCAGAAAGCGGAGCATGGCTGGATAAATCTTCCCGGTTTGGGATATTTAGTCCAAAA 1080
2119 CCAGAAAGCGGAGCATGGCTGGATAAATCTTCCCGGTTTGGGATATTTAGTCCAAAA 2178
1081 ACTTGAAGAGATACAGACAGACT 1105
2179 ACTTGAAGAGATACAGACAGATT 2203

RESULT 7
AX409637
LOCUS AX409637 13957 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2284 from Patent WO229103.
ACCESSION AX409637
VERSION AX409637.1 GI:21442342
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source
1..13957
/organism="Homo sapiens"
/mol_type="genomic DNA"

/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 13957;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTATGCTTACACAGAGCTGCTTATGTCACCACTCTGACCCCTACAGGAGCCATT 60
DB GAGTATGCTTACACAGAGCTGCTTATGTCACCACTCTGACCCCTACAGGAGCCATT 1158

QY 61 TCCTTACAGCATTTGGAAGCTCCTGGAACAAGTTCATTTGGCAGTTCATTTGATGAGAG 120
DB TCCTTACAGCATTTGGAAGCTCCTGGAACAAGTTCATTTGGCAGTTCATTTGATGAGAG 1218

QY 121 TGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAGATTTATCGTGCTCTTTC 180
DB TGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAGATTTATCGTGCTCTTTC 1278

QY 181 TGCTGAGACACATTTGGAAGCTCCTGGAACAAGTTCATTTGGAAGTTCATTTGGAAG 240
DB TGCTGAGACACATTTGGAAGCTCCTGGAACAAGTTCATTTGGAAGTTCATTTGGAAG 1338

QY 241 CCAGTTTACATCTACAGAGGCTGCTTATGTCACCACTCTGACCCCTACAGGAGCCATT 300
DB CCAGTTTACATCTACAGAGGCTGCTTATGTCACCACTCTGACCCCTACAGGAGCCATT 1398

QY 301 TAATATTTCTCAANTTGGGAAGTAAAGTTCATTTGGAACAAGTTCATTTGGAAGTAA 360
DB TAATATTTCTCAANTTGGGAAGTAAAGTTCATTTGGAACAAGTTCATTTGGAAGTAA 1458

QY 361 AACTGAAATCTACAGAGCAGATGATCTCTTAAATTCAGATGGAATGCTCAGGATAG 420
DB AACTGAAATCTACAGAGCAGATGATCTCTTAAATTCAGATGGAATGCTCAGGATAG 1518

QY 421 TAGCATGGAAGAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 480
DB TAGCATGGAAGAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 1578

QY 481 GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAGAAAGAAACAGGAAATGAGGAAGA 540
DB GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAGAAAGAAACAGGAAATGAGGAAGA 1638

QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTCAACACATTAAGTGCTTCA 600
DB GCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTCAACACATTAAGTGCTTCA 1698

QY 601 AGAAGATCTAGAACAGAACAGTCAAGGCTCAATCTCTCAGTCAATGCTGCTGATG 660
DB AGAAGATCTAGAACAGAACAGTCAAGGCTCAATCTCTCAGTCAATGCTGCTGATG 1758

QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGG 720
DB TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAAACAACCTTAAGGTATTGG 1818

QY 721 AGATCGATGGCAACATCTGTAGATGGAACAAGACCGCTGGTCTTTTACAGACAT 780
DB AGATCGATGGCAACATCTGTAGATGGAACAAGACCGCTGGTCTTTTACAGACAT 1878

QY 781 CCTTCTCAAAATGGCAACGCTCTTCTAGGAACAGTGCCTTTTATGTCATGCTTTTCA 840
DB CCTTCTCAAAATGGCAACGCTCTTCTAGGAACAGTGCCTTTTATGTCATGCTTTTCA 1938

QY 841 AAAAGAAGATCAGTGAACAAGATTTCAACAACCTGGCTTTTAAAGATCAAAATGAATGTT 900
DB AAAAGAAGATCAGTGAACAAGATTTCAACAACCTGGCTTTTAAAGATCAAAATGAATGTT 1998

QY 901 ATCAAGTCTTCAAAATCGGCGCTTTTAAAGCGGATCTAGAAAGAAAGCAATCCAT 960
DB ATCAAGTCTTCAAAATCGGCGCTTTTAAAGCGGATCTAGAAAGAAAGCAATCCAT 2058

QY 961 GGGCAAACTGATTTCACTCAAAACAAGATCTTTCTTCACTCAAGATTAAGTCAGTGAC 1020
DB GGGCAAACTGATTTCACTCAAAACAAGATCTTTCTTCACTCAAGATTAAGTCAGTGAC 2118

QY 1021 CCAGAAGACGGAAGCATGGCTGGATTAACCTTTGCCCGTGTGGGATTAATTTAGTCCAAA 1080
DB CCAGAAGACGGAAGCATGGCTGGATTAACCTTTGCCCGTGTGGGATTAATTTAGTCCAAA 2178

QY 1081 ACTTGAAAGAGTACAGCAGACT 1105
DB ACTTGAAAGAGTACAGCAGACT 2203

RESULT 8
AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
LOCATION/Qualifiers
1. 13957
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 13957;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTATGCTTACACAGAGCTGCTTATGTCACCACTCTGACCCCTACAGGAGCCATT 60
DB GAGTATGCTTACACAGAGCTGCTTATGTCACCACTCTGACCCCTACAGGAGCCATT 1158

QY 61 TCCTTACAGCATTTGGAAGCTCCTGGAACAAGTTCATTTGGCAGTTCATTTGATGAGAG 120
DB TCCTTACAGCATTTGGAAGCTCCTGGAACAAGTTCATTTGGCAGTTCATTTGATGAGAG 1218

QY 121 TGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAGATTTATCGTGCTCTTTC 180
DB TGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAGATTTATCGTGCTCTTTC 1278

QY 181 TGCTGAGACACATTTGGAAGCTCCTGGAACAAGTTCATTTGGAAGTTCATTTGGAAG 240
DB TGCTGAGACACATTTGGAAGCTCCTGGAACAAGTTCATTTGGAAGTTCATTTGGAAG 1338

QY 241 CCAGTTTACATCTACAGAGGCTGCTTATGTCACCACTCTGACCCCTACAGGAGCCATT 300
DB CCAGTTTACATCTACAGAGGCTGCTTATGTCACCACTCTGACCCCTACAGGAGCCATT 1398

QY 301 TAATATTTCTCAANTTGGGAAGTAAAGTTCATTTGGAACAAGTTCATTTGGAAGTAA 360
DB TAATATTTCTCAANTTGGGAAGTAAAGTTCATTTGGAACAAGTTCATTTGGAAGTAA 1458

QY 361 AACTGAAATCTACAGAGCAGATGATCTCTTAAATTCAGATGGAATGCTCAGGATAG 420
DB AACTGAAATCTACAGAGCAGATGATCTCTTAAATTCAGATGGAATGCTCAGGATAG 1518

QY 421 TAGCATGGAAGAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 480
DB TAGCATGGAAGAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 1578

ELTDLWLLSDQVKSQRVWVGLDLEINEMI KOKATMOOLEORRPOLEELI.TAAQNKLK
NKTSSOEASTI.LTDRIERTONQWDEVOHLONEROOLNEMLKDSQWMLBAKEEBOVL
GOARAKLESWEKGPYVDIAOKI.TETQLOAKDLQWQNTQVUDVADYAKLILROYSD
DTRKVMHITEN.NASRSHLKHRSEREALESRETHRELOQFPLDLKFLANLITEAETHA
NVLDQATRERLLEDKSGVKELMKQWDDUQGEIEAHTDVYHMLDENSOKILRSLEGGSD
DAVLLQRLDNNFKWSELRKKSINIRSHLEASSQWKLHLSLQELLWILQKXDEL
SRQAP.TGGDFPAQKQNDVHRAFKRELKTPKEPVINSTLETVRIFLTFEOLGLEKLYQ
EPRELPPERAAQNVTLKQAEVNTWEKLNLSADWOKI.DTEBLERLOLEQVATD
EIDLRLQAEVLKSGNQPVGLDILSLDHLKVKALRGEIAPLXENSHVNDLAROL
ETLGLQLSXPNLSTLHETQTTWDFHPKMETLYQSILADLNNVRFSAYRTAMKRLRLOKA
WERASPNKVPYIINHETQTTWDFHPKMETLYQSILADLNNVRFSAYRTAMKRLRLOKA
LCLDLSLSAADCADLDOHNLKNDQPMFDLIQINCLTIYDRLEHNNLNVPLUCVD
MCLNLLNVDYDGTGRIRVLSFKTGIISLCKAHLDEKRYLFLQVASTFCFOORRL
GMLIHDSIOI.PQLGVEAFSGSNTEPVSRCFQFANNKPEIEAALFLDWMRLPEQSM
VILPLVRHVAATAKHAQKACIKCEPIIGPRVSLKHFNVDICQSCFFPGRVAKGH
KVHVMVECTPTSTGSDVDFAKVLKNKFRKRYFAKHPRMGLYPVOTVLEGNMET
PVTLLNFMEVDASPNSSQPSRPSQAQILISLESERBELERILADLEBERNRLQAEY
DRUKQHEHKGUSPLPSPPEMPTSPSPRDAELIAEAKLLRQHKRLRLEARMQILDEH
NKQLESQHLRLQLLEPOQAEAKVNGTIVSSPSTLORSDSQPMRLLRVVSQSDSM
GEEEDLSPQDSTSTGLEEYNEQLNNSFPSSRGRTPGKPMREDTW"

BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 52.5%; Score 1103.4; DB 9; Length 13957;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GAGCTATGCTACACAGAGCTCTTATGTGACCACTCTGACCCCTACACGAGGCCATT 60
1099 GAGCTATGCTACACAGAGCTCTTATGTGACCACTCTGACCCCTACACGAGGCCATT 1158

61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAGTCAATTTGGCAGTTCAATGATGAGAG 120
1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAGTCAATTTGGCAGTTCAATGATGAGAG 1218

121 TGAAGTAACTCGACCGTATCAACAGCTTTAGAGAGATTTATCTGCGCTCTTTC 180
1219 TGAAGTAACTCGACCGTATCAACAGCTTTAGAGAGATTTATCTGCGCTCTTTC 1278

181 TGCTGAGGACATTTGCAAGACAGGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1279 TGCTGAGGACATTTGCAAGACAGGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338

241 CGAGTTTCACTACAGAGGATGATGATGATTTGACAGCCCATCAGGCGGGTGG 300
1339 CGAGTTTCACTACAGAGGATGATGATGATTTGACAGCCCATCAGGCGGGTGG 1398

301 TAATATTTCTACATTTGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAAGTGAAGA 360
1399 TAATATTTCTACATTTGGAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAAGTGAAGA 1458

361 AACTGAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 420
1459 AACTGAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 1518

421 TAGCATGGAAGAACAGCAATTTACATAGATTTTAAATGATCTCCAGAAATCAGAACT 480
1519 TAGCATGGAAGAACAGCAATTTACATAGATTTTAAATGATCTCCAGAAATCAGAACT 1578

481 GAAAGAGTTGAAATGACTGGCTTAACAAAAACAGAAAGAACAGGAAATGGAGGAAGA 540
1579 GAAAGAGTTGAAATGACTGGCTTAACAAAAACAGAAAGAACAGGAAATGGAGGAAGA 1638

541 GCCTCTTGGACCTGATCTGAAGACCTTAAACGCCAGTAGTACACATAGGTGCTTCA 600
1639 GCCTCTTGGACCTGATCTGAAGACCTTAAACGCCAGTAGTACACATAGGTGCTTCA 1698

601 AGAAGATCTAGAAACAGAAACAGTCAAGGTCAATTTCTCTCACTACATGCTGTGTAGT 660
1699 AGAAGATCTAGAAACAGAAACAGTCAAGGTCAATTTCTCTCACTACATGCTGTGTAGT 1758

661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 720

1759 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 1818
721 AGATCGATGGGCAACACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGACAT 780
1819 AGATCGATGGGCAACACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGACAT 1878
781 CCTTCTCAAAATGGCAACGCTTCTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTTACA 840
1879 CCTTCTCAAAATGGCAACGCTTCTACTGAAGAACAGTGCCTTTTATGTCATGGCTTTTACA 1938
841 AAAAGAGATGCGAGTGAACAGGATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
1939 AAAAGAGATGCGAGTGAACAGGATTCACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1998
901 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
1999 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 2058
961 GGGCAAACTGATTTCACTCAAAACAGATCTTCTTCAACTGAAGAAATAGTCACTGAC 1020
2059 GGGCAAACTGATTTCACTCAAAACAGATCTTCTTCAACTGAAGAAATAGTCACTGAC 2118
1021 CCAGAACAGCGAAGCATGCTGGATTAATTTGCCGCTGTGGGATAAATTTAGTCCAAA 1080
2119 CCAGAACAGCGAAGCATGCTGGATTAATTTGCCGCTGTGGGATAAATTTAGTCCAAA 2178
1081 ACTTGAAGAGATGACACAGACT 1105
2179 ACTTGAAGAGATGACACAGACT 2203

RESULT 10
AR220819 13977 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 60 from patent US 6426186.
ACCESSION AR220819
VERSION AR220819.1 GI:23327696
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13977)
AUTHORS Jones,K.A., Volkmut,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 60 30-JUL-2002;
FEATURES
Location/Qualifiers
1. 13977
/organism="unknown"
BASE COUNT 4596 a 2765 c 3120 g 3453 t 43 others
ORIGIN

Query Match 52.5%; Score 1103.4; DB 6; Length 13977;
Best Local Similarity 99.9%; Pred. No. 3.6e-251;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GAGCTATGCTACACAGAGCTGCTTATGTGACCACTCTGACCCCTACACGAGGCCATT 60
1099 GAGCTATGCTACACAGAGCTGCTTATGTGACCACTCTGACCCCTACACGAGGCCATT 1158
61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAGTCAATTTGGCAGTTCAATGATGAGAG 120
1159 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAGTCAATTTGGCAGTTCAATGATGAGAG 1218
121 TGAAGTAACTCGACCGTATCAACAGCTTTTAAAGAAAGTATTATCGTGGCTTCTTTC 180
1219 TGAAGTAACTCGACCGTATCAACAGCTTTTAAAGAAAGTATTATCGTGGCTTCTTTC 1278
181 TCCTGAGGACATTTGCAAGACAGGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
1279 TCCTGAGGACATTTGCAAGACAGGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
241 CCAGTTTCACTACAGGATCAGCAACTGCTGCTTTGGAAAGAACAACTTAAGGTATTGGG 300

Db	1339	CCAGTTTCATCTCATGAGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG	1398	
QY	301	TAATATTCATCAATTTGGAAGTAACTGATTTGGAACAGGAAATATATCAGAGATGAAGA	360	
Db	1399	TAATATTCATCAATTTGGAAGTAACTGATTTGGAACAGGAAATATATCAGAGATGAAGA	1458	
QY	361	AACTGAAGTACAAAGACAGATGAATCTCTAAATCAAGATGGGAATGCTTCAGGGTAGC	420	
Db	1459	AACTGAAGTACAAAGACAGATGAATCTCTAAATCAAGATGGGAATGCTTCAGGGTAGC	1518	
QY	421	TAGCATGGAATAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT	480	
Db	1519	TAGCATGGAATAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT	1578	
QY	481	GAAAGAGTTGAATGACTGCTTAACAAAACAGAAAGAAACAAGGAAATGGAAGGAAGA	540	
Db	1579	GAAAGAGTTGAATGACTGCTTAACAAAACAGAAAGAAACAAGGAAATGGAAGGAAGA	1638	
QY	541	GCCTCTTGACCTGATCTTGAAGACCTTAAAGACCTAAAGCCCAAGTACAAACAATAGGTCTTCA	600	
Db	1639	GCCTCTTGACCTGATCTTGAAGACCTTAAAGACCTAAAGCCCAAGTACAAACAATAGGTCTTCA	1698	
QY	601	AGAGATCTAGAACAGACAGTACAGGTCAATCTCTCACTCAGATGGTGGTAGT	660	
Db	1699	AGAGATCTAGAACAGACAGTACAGGTCAATCTCTCACTCAGATGGTGGTAGT	1758	
QY	661	TGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	720	
Db	1759	TGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG	1818	
QY	721	AGATCGATGGGCAACATCTGTAGATGACAGACAGCGTGGTCTTTTACAGAGCAT	780	
Db	1819	AGATCGATGGGCAACATCTGTAGATGACAGACAGCGTGGTCTTTTACAGAGCAT	1878	
QY	781	CCTTCTCAATGCGCAAGCTTACTGAAGAACAGTGCCTTTTGTGATGATGCTTTCAGA	840	
Db	1879	CCTTCTCAATGCGCAAGCTTACTGAAGAACAGTGCCTTTTGTGATGATGCTTTCAGA	1938	
QY	841	AAAGAGATGAGTGAACAGATTCACAACTGGCTTTAAAGATCAAAATGAATGTT	900	
Db	1939	AAAGAGATGAGTGAACAGATTCACAACTGGCTTTAAAGATCAAAATGAATGTT	1998	
QY	901	ATCAAGTCTTCAAAACTGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	960	
Db	1999	ATCAAGTCTTCAAAACTGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT	2058	
QY	961	GGCAAACTGTATTCATCAACAGATCTTTTCAACACTGAAGATAAGTCAAGTAC	1020	
Db	2059	GGCAAACTGTATTCATCAACAGATCTTTTCAACACTGAAGATAAGTCAAGTAC	2118	
QY	1021	CCAGAGACGGAGCATGCTGATCACTTCCCGGCTGCTGGGATATTTAGTCCAAAA	1080	
Db	2119	CCAGAGACGGAGCATGCTGATCACTTCCCGGCTGCTGGGATATTTAGTCCAAAA	2178	
QY	1081	ACTTTGAAAAGAGTACAGCAGACT	1105	
Db	2179	ACTTTGAAAAGAGTACAGCAGACT	2203	
RESULT 11				
HSDMDR				
LOCUS				
DEFINITION	Human mRNA for dystrophin.			
ACCESSION	X14298			
VERSION	X14298.1			
KEYWORDS	Dmd gene; Duchenne muscular dystrophy; dystrophin.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 12446) Rosenthal, A., Speer, A., Billwitz, H., Cross, G.S., Forrest, S.M. and			

Davies, K.E.
Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous
Nucleic Acids Res. 17 (13), 5391 (1989)
89345106
2668885
2 (bases 1 to 12446)
Rosenthal, A.
Direct Submission
Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften
der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Roesle
Str. 10, 1115 Berlin Buch, DDR
see also M18533 and M20250 for Dmd seqs.; discrepancies compared to
M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449,
3687, 4229, 4504, 5075, 5332, 5630 and 7194.
Location/Qualifiers
1. .12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X Chromosomal, Xp21."
/tissue_type="muscle"
/dev_stage="fetal and adult."
99. .11156
/note="dystrophin (AA 1 - 3685)"
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/translation="MLWEEVEDCYEREDVQKFTKWNQAQFKFGKHENLPSDL
QDRILLDTGLTGLPKERKSTRVHALNKNALRVLQNNVNDVNGISTDIDVG
NKHLLTGLWNTILHWQVNVKNIWAGLQPTNSEKILLSWVRQSTRYQKLVNINFT
TWSGLALNLIHSHRPDLFDWNSVQCSATQRLEHAFNARIYQLGTEKLHDPEDV
DTYTPDKILMYITSLFQVLPPQVYSIAIOSEVEMLPKPPKVTKEHLEPDKSGSSSL
QTVSLAQYERTSSPKPRFYSAYTAAYTTSTPTSPSPSSEHLEPDKSGSSSL
MESEVNDYQTALEEVLSWLSYASDTLQAQGEISINDVEYKQDTHHEGYMDLTAH
QVRGNIILQGLSKLITGKLSDETEVEQEQNNLSRWECURVASMESKLNHLRVLM
DIQVXKLENDLWTKTEERTKEEPLDPLEDKQVQOHKVLQEDLQEDQVRVN
SLTHVMVDESGBHATALEQKLVGLDRWANTCRWTEDEWLLQDILLKQWLKTE
ECLFSLNEXEDAVNKIHTTQFKDQNMESLQKLAVALKADLEKKQLQSKGLYSIK
QDLSTLKNKSYTKTEAWLDFARWNLVQKLEKTAQISOAVTTTPSTQITVTM
EVITPTITREQLVHQAQELPPPPKPKQITVSEIRKLDVDITELHSWITSEA
QVSPFPAIFRKEGNFSDUKENAIEREKAEKFTDKLQASRSAGALVEQWNGVNA
DSIQASELNSRWTFECQLLSERLNLWLEYQNNIIFATINQLQQLQEQMTTAAENWLKIQ
PTTSPPTAIKQKTKCDEVNRLSLQPOERLKIATLKEKQEQMFLDADVFVAF
TNHGFQVDFDQVAREKELQITFDLTPMRYQTSAGTARTWQOSETKLSIPOLSTVDY
EIMEQRLGLALQSSLOQSQGLYVLTIVKMSKAPSEISRYQSEFEIEGRWK
KLSSQVHCQKLEQMNKLRKIQNHIOTLKWMAEVDVFLKEWPALQDSILKQKL
KQRLVDSDITQLPSLNSWEGGQKIKNEAPEFASBLETELKELTQWDMHCOQVY
ARKEALGDLKTVSLQKLSRQKEMHWMTQABEYLERFEYKTPDELQKAFEMKRAK
EAAQKQKAVKLTITSVNSVIAQAPVQAQELKLETLITNYQWCTKNGCKTKLE
EVNACWHELLSYLEKANKVNLNEVEFKLTENIPGABEISEVLSLENLMSHSDNP
NQLRILOTLTQGVNDELINEELETFNSRWELHEEAVRQKLEQSIQSAQETNS
LHLIQSLNFIQKOLAAYTADKVAQMPQEAQKIQSDLTSHIEISLEMKHNOQKEA
AQVLSQDVAQKQLQDVSMKFLPKQKPAFQRLQESKMIIDEVKNHLPALLETKSV
QEVVQSLNHCNVLKLSSEKSEVEMVITKQEQIQVQKQENKELDERVATKLHY
NELGAKVTERKQKLEKLSRQKEMHWMTQABEYLERFEYKTPDELQKAFEMKRAK
VAMGKATQKIEKQKMETFDQVNDITKWIIOADTLDESEKKQPOKQEDVNLKRAELND
EWNLILLEYQKMETFDQVNDITKWIIOADTLDESEKKQPOKQEDVNLKRAELND
IRPKVDSITRQDQANLMAHGDHCKLVEPQISELNHFAAISHRIKNGKASIPKLELE
QPNSDIKLLEPLEAEIQGVNLKEDFDNEDNEGTVELLQDNLQORLIDIEKKL
KSEIKIKQQLLOTKNALKOLRSORPKKLEI SHOWYOYKQADDLKCLDIEKKL
ASUPEPRDKRKEIDREILQKKEELNVAARGAGLSEDAAMAVEPTQIQISKRWRE
IESKFAQRLNFAQIRVRETMEDMLEISYVSTYLTETHVTSQALVEVQ
LNAPDLCAKDFEDFKBEESKNISQSQSSGRDIIHKKTAALQALQAFVETVKL
QEALSQDFOKENVKRYQKQRFDRSVKRRFYDIKIFNQWTEAEQFURKQI
PENWEHAKYKWLKELQDQIGQRTVVRVLTNATGEEIIQSSKTASIIQKLSGLN
RWQVCKQSLDRKRKEOKNLSBFORDLNELFVLEADNATISPLBPGKEQQLKE
KLSEVLLUEBELPQROGLKQNLTEGTVLVSAPISPEODKLENKQNLQWIKVS
RALPEKQGEAEQIKQGLQKLELLEQLNELLMLSPRNQLYVNPQNGEPFD
KQVTEAVQAKQPDVEILSKGQHLKEKPAQPVKRLKLEDSSEKHAVNRLLOELRA
KQDPLAPGLTTIGASPTQVTLVTPQVTKTAKISLEMPSSMLSEVLPALDFNRWAT
ELTDWLSLLDQVTKSQSRVMVMDLEDINEMI IKQATMDQLBQRPRPQLELITAAQNLK

BASE COUNT 4135 a 2524 c 2876 g 2911 t
ORIGIN
Query Match 52.4%; Score 1100.2; DB 9; Length 12446;
Best Local Similarity 99.7%; Pred. No. 2.1e-250;
Matches 1102; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
2y 1 GAGCTATGCTACACAGAGCTGCTTATGTCACCACTCTGACCTCTGACCTCTACAGGAGCCATT 60
db 989 GAGCTATGCTACACAGAGCTGCTTATGTCACCACTCTGACCTCTGACCTCTACAGGAGCCATT 1048
2y 61 TCCTTCACAGCATTTGGAAGCTCTCTGACAGCAAGTCAATTTGGCAGTTTCAATGATGAGAG 120
db 1049 TCCTTCACAGCATTTGGAAGCTCTCTGACAGCAAGTCAATTTGGCAGTTTCAATGATGAGAG 1108
2y 121 TGAAGTAACTCGAGCGGTTATCAACAGCTTTAGAGAGATTTATCTGCTCTCTTTC 180
db 1109 TGAAGTAACTCGAGCGGTTATCAACAGCTTTAGAGAGATTTATCTGCTCTCTTTC 1168
2y 181 TGCTGAGGACACATTTGGAAGCTCTCTGACAGCAAGTCAATTTGGCAGTTTCAATGATGAGAG 240
db 1169 TGCTGAGGACACATTTGGAAGCTCTCTGACAGCAAGTCAATTTGGCAGTTTCAATGATGAGAG 1228
2y 241 CCAGTTTCACTCATGAGGAGTACATGATGATTTGACAGCCCATCAGGCGCGGTTGG 300
db 1229 CCAGTTTCACTCATGAGGAGTACATGATGATTTGACAGCCCATCAGGCGCGGTTGG 1288
2y 301 TAATATTTCAATTTGGAAGTAAAGTGTGAGTGAACAGGAAATTTATCAGAGATGAAGA 360
db 1289 TAATATTTCAATTTGGAAGTAAAGTGTGAGTGAACAGGAAATTTATCAGAGATGAAGA 1348
2y 361 AACTGAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGAATGCTCAGGCTAGC 420
db 1349 AACTGAGTACAGAGCAGATGATCTCTTAATTTCAAGATGGAATGCTCAGGCTAGC 1408
2y 421 TAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAAT 480
db 1409 TAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAAAT 1468
2y 481 GAAAGATTTGAATGATGCTGCTTACAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
db 1469 GAAAGATTTGAATGATGCTGCTTACAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1528
2y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACCAACATTAAGTGCTTCA 600
db 1529 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACCAACATTAAGTGCTTCA 1588
2y 601 AGAAGATTTAGAAACAGAAACAGTTCAGGCTCAATTTCTCACTCAATGCTGCTGATG 660
db 1589 AGAAGATTTAGAAACAGAAACAGTTCAGGCTCAATTTCTCACTCAATGCTGCTGATG 1648
2y 661 TGATGATCTAGTGGATCAGCAAGCTGCTGCTTTGGAGAGCAACTTAAGGTATTGGG 720

Db 1649 TGATGATCTAGTGGATCAGCAAGCTGCTGCTTTGGAAGAACAACTTAAGGTATTGGG 1708
Qy 721 AGATCGATGGGCAAAACATCTGTATGATGAGACAGAGACCGCTGGTCTTTTACAGACAT 780
Db 1709 AGATCGATGGGCAAAACATCTGTATGATGAGACAGAGACCGCTGGTCTTTTACAGACAT 1768
Qy 781 CCTTCTCAATGCAAGCTGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTACAG 840
Db 1769 CCTGCTCAATGCAAGCTGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTTACAG 1828
Qy 841 AAAAGAGATGCGATGAGCAAGATTTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 900
Db 1829 AAAAGAGATGCGATGAGCAAGATTTACACAACTGGCTTTTAAAGATCAAAATGAAATGTT 1888
Qy 901 ATCAAGTCTTCAAAACTGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 960
Db 1889 ATCAAGTCTTCAAAACTGCGCTTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1948
Qy 961 GGGCAAACTGATTTCACTCAAAACAGATCTTTTCAACACTGAAGAAATAGTCAGTGAC 1020
Db 1949 GGGCAAACTGATTTCACTCAAAACAGATCTTTTCAACACTGAAGAAATAGTCAGTGAC 2008
Qy 1021 CCAGAGACGGAAGATGCTGGATTAATCTTTGCCCGTGTGGATTAATTTAGTCCAAA 1080
Db 2009 CCAGAGACGGAAGATGCTGGATTAATCTTTGCCCGTGTGGATTAATTTAGTCCAAA 2068
Qy 1081 ACTTGAAGAGATGACACAGACT 1105
Db 2069 ACTTGAAGAGATGACACAGACT 2093
RESULT 12
AX538620
LOCUS AX538620 5339 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 40 from Patent WO0229056.
ACCESSION AX538620
VERSION AX538620.1 GI:25271166
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 40 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source
1..5339
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1638 a 1191 c 1187 g 1323 t
ORIGIN
Query Match 47.5%; Score 997; DB 6; Length 5339;
Best Local Similarity 100.0%; Pred. No. 6.9e-226;
Matches 997; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1105 TCATAGATTACTGCAACAGTTCCCTCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1164
Db 1546 TCATAGATTACTGCAACAGTTCCCTCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGA 1605
Qy 1165 AGCTGAAACAACTCCCAATGCTTACAGGATGCTACCCGTAAGGAAAGGCTCTTAGAGA 1224
Db 1606 AGCTGAAACAACTCCCAATGCTTACAGGATGCTACCCGTAAGGAAAGGCTCTTAGAGA 1665
Qy 1225 CTCCAGGAGTAAAGAGCTGATGAACATGCTCCAGAGCTCCCAAGGTGAATTTGAAGC 1284
Db 1666 CTCCAGGAGTAAAGAGCTGATGAACATGCTCCAGAGCTCCCAAGGTGAATTTGAAGC 1725
Qy 1285 TCACACAGATGTTTATCACAACCTGGATGAAAACAGGCCAAAATCTCTGAGATCCCTGGA 1344

1726 TCACACAGATGTTTATCACAACTGGATGAAACAGCCAAAAATCTGAGATCCCTGGA 1785
1345 AGTTTCCGATGATGAGTCCCTGTTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAG 1404
1786 AGTTTCCGATGATGAGTCCCTGTTTACAAAGACGTTGGATAACATGAACTTCAAGTGGAG 1845
1405 TGAATCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGAAGCCAGTCTTGACCAAGT 1464
1846 TGAATCTCGGAAAAAGTCTCTCAACATTAGTCCCATTTTGAAGCCAGTCTTGACCAAGT 1505
1465 GAAGGCTGTGACCTTTCTCTGACGAACTTCTGTTGGTGGTCAAGTCAAGTGAATGTA 1524
1906 GAAGGCTGTGACCTTTCTCTGACGAACTTCTGTTGGTGGTCAAGTCAAGTGAATGTA 1965
1525 ATTAAGCCGGCAGGACCTATTGAGGCGACCTTTCCAGCAGTCTCAGAAAGCAGAAACGATGT 1584
1966 ATTAAGCCGGCAGGACCTATTGAGGCGACCTTTCCAGCAGTCTCAGAAAGCAGAAACGATGT 2025
1585 ACATAGGGCTTCAAGAGGAAATTGAAGAACTTAAGAACTTGAAGAACTTGAAGAACTTGA 1644
2026 ACATAGGGCTTCAAGAGGAAATTGAAGAACTTGAAGAACTTGAAGAACTTGAAGAACTTGA 2085
1645 GACTGTAGCAATATTCTGACAGAGCAGCTTTTGAAGAACTTGAAGAACTTGAAGAACTTGA 1704
2086 GACTGTAGCAATATTCTGACAGAGCAGCTTTTGAAGAACTTGAAGAACTTGAAGAACTTGA 2145
1705 GCCAGAGAGTGCCTCTGAGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGCA 1764
2146 GCCAGAGAGTGCCTCTGAGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGCA 2205
1765 GCGTGAAGAGTCAATACCTGAGTGGGAAATTTGAAGAACTTGAAGAACTTGAAGAACTTGA 1824
2206 GCGTGAAGAGTCAATACCTGAGTGGGAAATTTGAAGAACTTGAAGAACTTGAAGAACTTGA 2265
1825 AAAAAATAGATGAGACCTTTGAAGAACTTGAAGAACTTGAAGAACTTGAAGAACTTGAAG 1884
2266 AAAAAATAGATGAGACCTTTGAAGAACTTGAAGAACTTGAAGAACTTGAAGAACTTGAAG 2325
1885 CTTCAAGTGCCTCAAGTCTGAGTGAATCAAGAGTCTGAGAGCCAGTGGGCGATCTCCT 1944
2326 CTTCAAGTGCCTCAAGTCTGAGTGAATCAAGAGTCTGAGAGCCAGTGGGCGATCTCCT 2385
1945 CATTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCAGAGGCAAAATTCGCCC 2004
2386 CATTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCAGAGGCAAAATTCGCCC 2445
2005 TCTGAAGAGAACTGAGCACTCAATGACCTTGAAGTCTGAGAGGCAAAATTCGCCC 2064
2446 TCTGAAGAGAACTGAGCACTCAATGACCTTGAAGTCTGAGAGGCAAAATTCGCCC 2505
2065 TCAGCTCTACCGTATTAACCTCAGCACTCTGGAAGAC 2101
2506 TCAGCTCTACCGTATTAACCTCAGCACTCTGGAAGAC 2542

RESULT 13
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1..5462
/organism="synthetic construct"

/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1668 a 1225 c 1212 g 1357 t
ORIGIN
Query Match 47.4%; Score 996; DB 6; Length 5462;
Best Local Similarity 100.0%; Pred. No. 1.2e-225; Indels 0; Gaps 0;
Matches 996; Conservative 0; Mismatches 0
QY 1106 CATAGATTACTCAACAGATTCCTCCCTGGACCTGGAAAAAGTTTCTTTCCTGCTGTACAGAA 1165
Db 1670 CATAGATTACTCAACAGATTCCTCCCTGGACCTGGAAAAAGTTTCTTTCCTGCTGTACAGAA 1729
QY 1166 GGTGAAAAAATCTGCAATGTCTCAGAGATGCTACCCCTTAAGAAAGGCTCTCTAGAGAC 1225
Db 1730 GGTGAAAAAATCTGCAATGTCTCAGAGATGCTACCCCTTAAGAAAGGCTCTCTAGAGAC 1789
QY 1226 TCCAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAAGGTGAAATTGAAGCT 1285
Db 1790 TCCAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAAGGTGAAATTGAAGCT 1849
QY 1286 CACACAGATGTTTATCAACAACTGATGAAACAAAGCAAGCAAAAAATCTCTGAGATCCCTGGAA 1345
Db 1850 CACACAGATGTTTATCAACAACTGATGAAACAAAGCAAGCAAAAAATCTCTGAGATCCCTGGAA 1909
QY 1346 GGTTCGATGATGCACTCTCTGACAGAGACCTTTGGATTAACATGAATCTTCAAGTGGAGT 1405
Db 1910 GGTTCGATGATGCACTCTCTGACAGAGACCTTTGGATTAACATGAATCTTCAAGTGGAGT 1969
QY 1406 GAACTTCGAAAAAGTCTCTCAACATTTAGGTCCATTTGGAAAGCCAGTTCAGACCAAGTGG 1465
Db 1970 GAACTTCGAAAAAGTCTCTCAACATTTAGGTCCATTTGGAAAGCCAGTTCAGACCAAGTGG 2029
QY 1466 AACGCTCTGCACTTTCTCTGACAGAGACCTTTGGTGTGCTACAGCTGAAAGATGATGAA 1525
Db 2030 AACGCTCTGCACTTTCTCTGACAGAGACCTTTGGTGTGCTACAGCTGAAAGATGATGAA 2089
QY 1526 TTAAGCCGGCAGGACCTATTGAGAGCCAGCTTTTCCAGCAGTTCAGAAAGCAGAACTGTA 1585
Db 2090 TTAAGCCGGCAGGACCTATTGAGAGCCAGCTTTTCCAGCAGTTCAGAAAGCAGAACTGTA 2149
QY 1586 CATAGGGCTTCAAGAGGAAATTTGAAGAACTTGAAGAACTTGAAGAACTTGAAGAACTTGA 1645
Db 2150 CATAGGGCTTCAAGAGGAAATTTGAAGAACTTGAAGAACTTGAAGAACTTGAAGAACTTGA 2209
QY 1646 ACTGTACGAATATTTCTGACAGAGACCTTTTGGAAAGGACTAGAGAACTTCTACCAAGGAG 1705
Db 2210 ACTGTACGAATATTTCTGACAGAGACCTTTTGGAAAGGACTAGAGAACTTCTACCAAGGAG 2269
QY 1706 CCCAGAGAGTGCCTCTCTGAGGAGAGGCCAGAAATGTCTCTGCGCTTCTACGAAAGCAG 1765
Db 2270 CCCAGAGAGTGCCTCTCTGAGGAGAGGCCAGAAATGTCTCTGCGCTTCTACGAAAGCAG 2329
QY 1766 GCTGAGAGGCTCAATCTGAGTGGAAAAATTTGAACTCTGCACTCCGCTGACTGGCAGAGA 1825
Db 2330 GCTGAGAGGCTCAATCTGAGTGGAAAAATTTGAACTCTGCACTCCGCTGACTGGCAGAGA 2389
QY 1826 AAAATAGATGAGACCTTGAAGAACTTCAAGAGCCACCGATAGCTGGAC 1885
Db 2390 AAAATAGATGAGACCTTGAAGAACTTCAAGAGCCACCGATAGCTGGAC 2449
QY 1886 CTCAGCTGGCCAGCTGAGGTGATCAAGGGATCTGGCAGACCTGGGCGGATCTCCTC 1945
Db 2450 CTCAGCTGGCCAGCTGAGGTGATCAAGGGATCTGGCAGACCTGGGCGGATCTCCTC 2509
QY 1946 ATTCACTCTCTCCAAAGATCACTCGAGAAAGTCAAGCACTTCGAGAGAAATTCGCGCT 2005
Db 2510 ATTCACTCTCTCCAAAGATCACTCGAGAAAGTCAAGCACTTCGAGAGAAATTCGCGCT 2569
QY 2006 CTGAAAGAGAACTGAGGACCGATCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCAAT 2065
Db 2570 CTGAAAGAGAACTGAGGACCGATCAATGACCTTGTCTGCCAGCTTACCACTTTGGGCAAT 2629

QY 2066 CAGCTCTCACCGGTATACCTCAGCACTCTGGAAGAC 2101
|||||
Db 2630 CAGCTCTCACCGGTATACCTCAGCACTCTGGAAGAC 2665

RESULT 14

AF070485 13887 bp mRNA linear MAM 09-DEC-1998
Canis familiaris dystrophin mRNA, complete cds.

AF070485
AF070485.1 GI:3982750

Canis familiaris (dog)

Canis familiaris

Canis familiaris

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Canis familiaris

1. (bases 1 to 13887)

Carville, K.S., Mann, C.J., Scharzberg, S.J. and Wilton, S.D.

Direct Submission

Submitted (04-JUN-1998) ANRI, Pathology, University of Western

Australia, Verdun Street, Nedlands, WA 6018, Australia

Location/Qualifiers

1. 13887

/organism="Canis familiaris"

/mol_type="mRNA"

/strain="breed Golden retriever"

/db_xref="taxon:9615"

<1. 1387

290. 11332

/codon_start=1

/product="dystrophin"

/protein_id="AAC83646.1"

/db_xref="GI:3982751"

/translation="MLWNEVEDEVCYREDVQKTKFTKWNVAQSFPGKHIEHNPESDL

QDGRLLDLEGGTQKPKKSTRVHALLNNVKNALRVLQXNVDLVNIGSDIVDQ

NKHLTLGLINLTHSRPDLFDNNSVVCQCSATRELEHAFNIAKVLQGLG:EKLLDPEDV

ATYPPDKSLMYTSLFOVLPOQVSEIQAQVEMLPSPQVTRREHFOIHEOMYSQ

QIVSVLAQGERAPSPKPKFSAVYATQAYVTTSPTSPSLPSOHLPTEDPKSGRS

LTETENLDSYQALBEVSLWSAEQAQAQGEINSDVBEVEKQPHHEGYMMLTS

HQGRVGNQLGSLQIGTQKLSDEDETEVOENLNLNSWECURVASMKEQNSLHKVL

MDLQNGQLKELNMTKTEERTKMEKPELPGDIEDLKEQVQKHVLQDELEQEVRY

NSLTTHMVVDESGSHATAAEQLKVLGRWANI CRWTEDEWVLLQDILLKWRFT

BEOLFSALSEKEDAVNKHTTGFKDQSEVLSNQLAVLTDLEKKQTMDKLCSL

NODLSALKNVVAHWKWEALDFAORWNLVQLEKSSAQISOAVTTTTPSLTQTV

MEIVWVTRHILVHQAQELPPFPQKRLIIVDSERKRLDVIDITELHSHWITRE

AVLOSPEFALYKREGEVDLKEKYNALEREKAEKFRKQDASRAQALVEQWNVGN

ADSIKQASEQINERWIEFCOLLSERLNWLLEYQNNITFTYNOQLQDEQMTTAAENLKT

QPTTTSPTAKSLQIKCKEINRLSALQPIERLXIQSIALKEKQGMFLDADFA

FTNHNOVFADVAQAREKELQTI FDSLPPMRYQETMTSLTWTIQSSTKLSIPQVTV

YDIMEQRLGLQALQSSLOQNGNLVLTSTVKEMSKAPLSDISRYQSEPEEIEGR

WKLSSOLVHCKLSEOMAKRLKONHILKAKWITVDVFLKBEWPAIGDSEILKR

QLKQRLVNDIQTIQSLNSVNEGQKMKNEAPEFAGLETELELNTQDYMCRQ

VIARKEALKGGLDKTQSLQSLSEMHMTQAEELDERDFETKTPDELQTAVERNKR

AKEEAQKQKAKLLTSVNSVTAQAPPAQEAALKEKLEPLTNTYQMLCTRNGKCT

LEEVAWCHELLSYLEKANKLSEVEYKLTNTENISGGAETAEVLDSLENLMQSHED

NPNOIRLAOTLTQVHDELINLELETNSRWRELEHAEVRRQKLEQSIQSAQIE

KSULHIOESLSSIDKQALAAVADDAQAQPEAKIQSLDTSHEISLEEMKHNOGK

ETAQRVLSQIDVAQKLQDYSMKFRLQKPAEFQBLQSKMLDLDEVKHLPALETKS

VEQVVSQNLHCVNLTSLSSEVSEVEMVITKGRQIVQKQNTENPEKLDERTALKL

HYNELGAKVTERQKLEKLSLRMKRKNWALTTEALTDMLTKRSVGEPMPSNLD

SEVANGATQKIEIKQVHLSVTEVEALTVLGKEMLVEDKSLNSNNTAVTSR

AEELWNLLEYKHMTFDQNVYITNWIQADALLDESEKKPOQKEDILKELKAM

NDIPKVSITRDQANLANMGDHCRCRVBPKESELNHPAALSHRIKTKASIPKE

LEQNSDIQLLEAEIIOGVNLKEEDFKMDSENEGTVKELLORGNLQORLTID

KRKEEBIKIQOOLLQTHNALKOLRSORRKALEISHQYQKROADDLLKCLDLEK

KLASIPRDERKIKEDIRELOKKEELNVAVRQAGSLSEDGAAMAVEPTQOLSKEK

REIESKFAQFRRLNAQTHVEESVVMTEDMPLSISYVFTYLTETITVYSQALSEV

KLQALSRDLQFQWRVNMVQKQRFDRSVKRWFRIDYIHNKTKAALHSAFPAERA

QIPENWEHAKYKWLKELQDQIGQSVRVNLNATGEEITQSSKTDASITQKGLSL

NLRWQECVKQLAERKKEQLKSLSEKQNFVNLVEEADNVANIPLQNEQOOL

KEKLEQVKLLAEELPRQGLIKOLNETGGTGLVLSAPLSPEEQDKLENKLIKQTNLQMIK
VSNLPEKQBEIEBAHYKDLQLEEQNLHLLMLSPIRNQLEIYNQNTQGPFDIKIEIE
VAVQAKQPVGEGISKGQHLYPEKPAQPKAKLEDLSLSDWKVVTQLQLQRAWTKP
APGLTTVRAPPSQTVTLVTPQAKETATSKLEMPSSLLLEVPALAFNRAMTETLTDW
LSULDRVRSQRMVWGLEDINEMI KOKATLOLQORRQLEELITAAQNLKNTSN
OEARTIITIERIQSOMDEVQSHLNRRLQLEMLKDSQWLEAKEAEQVILQQAARA
KLESWEAPYTVDAIOKKITETQKADIRQOINVDVANDLAKLLRDYSADDTKV
HMITENINASWASHIKLSREAALETHRLQOFLDLEKFLAWITEATTANVLQD
ATHKERLEDKGVRLMKQWOLQGEIEAHTDIYHNLDSQKRLHLSQELLVWLQKDELSRQAP
QRLIDNNFKWSLRKSLNIRSHLEASDQWRLHLSQELLVWLQKDELSRQAP
IGGDFPAVQKQNDVHRAFKRELKTEPFVIMSTLETVRIIFLQEPLEGLXLYQEPREL
PPEARAQNTVTLRLKQAAEVNTQWELNVHNSADWQRKIDAEALRLQLEQATDELDK
LQAEVTKGSQWQVGLLIDLSODHLEKVKALRGETPLKENVSYYNDIAROLTLGTI
QLSPYNLTLEDLNTRWKLLQVALEDRIQLEHARDFGSPASCHFSVQSGPWERAI
SPNKPYIYNHETQTCWHDPRFELYQSLADLNVRFSAYRTAMKRLRLQKALCLDL
LSLSAACDALDQHNKQNDQMDILQVINCITTIYDRLEQEHNNLVNVPUCVMDCLNH
LLANYDRTGRTGIRVLSFKTGISLCAHLEDKRYLFKQVASTGSGVLRVQVMDCLNH
DSIOIPQLGEVASFGSNIPEPSVRCFQANNKPEIEALFLDMWLEPQSMVWLPV
LHRVAAATAKQAKCNICEKPIIGFVRSIAKFNVDICQSCFFSFRVAKGHMHP
MVEXCTPTTSGEDVRDFAKVLKNKPKTKRYFAKHPRMGYPVQTVLEGDNMETPTVLI
NFPVDSAPASSQLSHDDTHSRLEHVASLKKWENSGSYLNDISIPNESIDDEHLL
IQHWRSLNQESPLSQSPSPAQILISLESERGERILADLGRNRLNQAEDRLKQ
QHEKGLSPLEPPEMPTSPQSRDAELIAEAKLLRQHKRLRLEARNQILEDHNRKQLE
SOHLRQLLEQQAQAEKVNQTTVSSPSTLSQSDSSQPMLLRVVSGTSSESMEGEEDL
LSPQDTSGLLEEVMEQLNHSFPSSRGRNTPFGKPMREDTW"

3'UTR

repeat_region

/rpt_type=tandem

/rpt_unit=ca

BASE COUNT 4455 a 2906 c 3164 g 3362 t

ORIGIN

Query Match 44.2%; Score 929; DB 4; Length 13887;

Best Local Similarity 90.0%; Pred. No. 9.8e-210;

Mismatches 995; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 1 GAGTATGCTTACACAGAGCTGCTTATGTCACACCTCTGACCCCTACAGGACCCCACT 60

Db 1183 GAGTATGCTTACACAGAGCTGCTTATGTCACACCTCTGACCCCTACAGGACCCCACT 1242

QY 61 TCCTTTCAGACATTTGGAAGCTCCTGGAAGCAAGTCAATTTGGCAGTTCAITGATGGAGAG 120

Db 1243 TCCTTTCAGACATTTGGAAGCTCCTGGAAGCAAGTCAATTTGGCAGTTCAITGATGGAGAG 1302

QY 121 TGAAGTAAACCTGACCGTTATCAAAACAGCTTTTGAAGAAGTATATCGTGGCTCTTTTC 180

Db 1303 CGAAGCAACCTGACAGCTTATCAAAACAGCTTTTGAAGAAGTATCTCTGCTGGCTCTTTTC 1362

QY 181 TCCTTTCAGACATTTGGAAGCTCCTGGAAGCAAGTCAATTTGGCAGTTCAITGATGGAGAG 240

Db 1363 AGCTGAGGATGCACTGCAAGCCCAAGGAGAGATTTCAATGATGTCGAAGAGTGAAGA 1422

QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCCCGGGTTGG 300

Db 1423 ACAATTCATCTCATGAGGGATATATGATGCACTTGCATCATCCATCAGGACGGGTGG 1482

QY 301 TAAATTTCTCAATTTGGGAAGTAACTGATTTGGAAGCAAGGAAATATCAGAGATGAGA 360

Db 1483 TAAATTTCTCAATTTGGGAAGTAACTGATTTGGAAGCAAGGAAATATCAGAGATGAGA 1542

QY 361 AACTGAAGTCAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420

Db 1543 AACCAAGTGGAGAACAAATCTCTCAATTCAGATGGGAATGCCTCAGGGTAGC 1602

QY 421 TAGCATGAAAAACAAACAAATTTATAGAGTTTTATAGATCTCAGAAATCAGAACT 480

Db 1603 TAGCATGAAAAACAAACAAATTTATAGAGTTTTATAGATCTCAGAAATCAGAACT 1662

QY 481 GAAAGAGTTGATGATCGCTGCTTAAACAAACAGAGAGAGAAACAAAGGAAATGAGGAGAGA 540

Db 1663 GAAAGAGTTGATGATCGCTGCTTAAACAAACAGAGAGAGAGAAACAAAGGAAATGAGGAGAGA 1722

QY 541 GCCTCTTGACCTGATCTTGAAGACCTTAAACAGCCCAAGTACAAACAAATAGGTCCTTCA 600

[illegible]

```

RESULT 15
E30220 LOCUS linear PAT 18-JUN-2001
DEFINITION Shortened dystrophin.
ACCESSION E30220
VERSION E30220.1 GI:13017027
KEYWORDS JP 199318467-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4402)
AUTHORS Sinichi,T.
TITLE Shortened dystrophin
JOURNAL Patent: JP 199318467-A 3 24-NOV-1999;
SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY
OS Unidentified
PN JP 199318467-A/3
PD 24-NOV-1999
PF 08-MAY-1998 JP 1998142134
PR
PI SINICHI TAKEDA
PC C12N15/09,A61K48/00,C12M15/00
CC Strandedness: Both;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4402 /organism='Unidentified'.
FT Location/Qualifiers
1..4402 /organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'
BASE COUNT 1339 a _984 c 1010 g 1069 t
FEATURES source

```

ORIGIN	Query Match	43.4%;	Score 911.6;	DB 6;	Length 4402;
	Best Local Similarity	99.6%;	Pred. No. 1.3e-205;		
	Matches 914;	Conservative	0;	Mismatches 4;	Indels 0; Gaps 0;
QY	1	GAGCTATGCTACACACAGCGTGCCTTATGCTACACACCTCTGACCCCTACCGAGCCCAATT	60		
Db	1095	GAGCTATGCTACACACAGCGTGCCTTATGCTACACACCTCTGACCCCTACCGAGCCCAATT	1154		
QY	61	TCGTTTCACAGCATTTGGAGCTCCTCGAAGCAAGTCAATTTGGCAGTTTCATTGATGGAGAG	120		
Db	1155	TCGTTTCACAGCATTTGGAGCTCCTCGAAGCAAGTCAATTTGGCAGTTTCATTGATGGAGAG	1214		
QY	121	TGAAGTAAACCTGGACCGGTTATCAACACAGCTTTAGAAAGATATTATCGTGGCTCTCTTTC	180		
Db	1215	TGAGTAAACCTGGACCGGTTATCAACACAGCTTTAGAAAGATATTATCGTGGCTCTCTTTC	1274		
QY	181	TGCTGAGGACATNTTCAAGCACCAAGGAGAGATTTCTAATGATGTGGAGTGTGTGAAGA	240		
Db	1275	TGCTGAGGACATNTTCAAGCACCAAGGAGAGATTTCTAATGATGTGGAGTGTGTGAAGA	1334		
QY	241	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG	300		
Db	1335	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG	1394		
QY	301	TAAATTTTCAATTTGGGAAGTAAAGTGTATGGACAGGAAATTTATCAGAACTGAAGA	360		
Db	1395	TAAATTTTCAATTTGGGAAGTAAAGTGTATGGACAGGAAATTTATCAGAACTGAAGA	1454		
QY	361	AACCTGAAGTACAAGACAGATGAATCTCTCTAAATTCAGATGGGAATGCTCTCAGGGTAGC	420		
Db	1455	AACCTGAAGTACAAGACAGATGAATCTCTCTAAATTCAGATGGGAATGCTCTCAGGGTAGC	1514		
QY	421	TAGCATGGAAAAACAAAGCAATTTACATAGATGTTTAAATGGATCTCCAGAACTCAGAACT	480		
Db	1515	TAGCATGGAAAAACAAAGCAATTTACATAGATGTTTAAATGGATCTCCAGAACTCAGAACT	1574		
QY	481	GAAAGAGTTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAATGGAGGAAGA	540		
Db	1575	GAAAGAGTTTGAATGACTGGCTTAACAAAAACAGAAAGAAAGAAATGGAGGAAGA	1634		
QY	541	GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACACACATAAAGTGCTTCA	600		
Db	1635	GCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACACACATAAAGTGCTTCA	1694		
QY	601	AGAAGATCTAGAACAAAGAACAGTCAAGGTCMAATCTCTCACTCACATGTTGTTGTTAGT	660		
Db	1695	AGAAGATCTAGAACAAAGAACAGTCAAGGTCMAATCTCTCACTCACATGTTGTTGTTAGT	1754		
QY	661	TGATGAATCTAGTGGAGATCACCCAACTGCTGCTTTGGAGAACACATTAAAGTATTTGGG	720		
Db	1755	TGATGAATCTAGTGGAGATCACCCAACTGCTGCTTTGGAGAACACATTAAAGTATTTGGG	1814		
QY	721	AGATCGATGGGCAAAACATCTGTAGATGGAAGAAGACCGCTGGGTTCTTTTACAAGACAT	780		
Db	1815	AGATCGATGGGCAAAACATCTGTAGATGGAAGAAGACCGCTGGGTTCTTTTACAAGACAT	1874		
QY	781	CCCTCTCAAAATGGCAACGCTCTTACTAGAAACAGTGGCTTTTATAGTCATGGCTTTTCAGA	840		
Db	1875	CCCTCTCAAAATGGCAACGCTCTTACTAGAAACAGTGGCTTTTATAGTCATGGCTTTTCAGA	1934		
QY	841	AAAAGAGATGTCAGTGAACAAAGATTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT	900		
Db	1935	AAAAGAGATGTCAGTGAACAAAGATTCACAACTGGCTTTTAAAGATCAAAATGAAATGTT	1994		
QY	901	ATCAAGTCTTCAAAAACCT 918			
Db	1995	ATCAAGTCTTCAAAAACCT 2012			

Search completed: February 1, 2004, 23:50:54
Job time : 7511.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:14:55 ; Search time 4634.58 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100

Perfect score: 1301

Sequence: 1 cgactttccagcagttcaga.....ccggagcgtctccaatttg 1301

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1301	100.0	5339	6	AX538620 Sequence
2	1301	100.0	5462	6	AX538621 Sequence
3	1301	100.0	5952	6	AX304538 Sequence
4	1301	100.0	5952	6	AX114289 Sequence
5	1301	100.0	8689	6	AX538622 Sequence
6	1301	100.0	11443	6	AX538624 Sequence
7	1301	100.0	12057	6	AX538627 Sequence
8	1301	100.0	12446	9	HSMDMR
9	1301	100.0	13957	6	AX409637 Sequence
10	1301	100.0	13957	6	AX538581 Sequence
11	1301	100.0	13957	9	HUMDYS
12	1290	99.2	13977	6	AR220819 Sequence
13	1154.4	88.7	13887	4	AF070485
14	1135.2	87.3	13815	6	AX306153
15	1135.2	87.3	13815	6	AX538582 Sequence
16	1135.2	87.3	13815	10	MUSDYSA
17	1135.2	87.3	19307	6	AR093392 Sequence
18	1135.2	87.3	19307	6	AR142592 Sequence
19	1085.2	83.4	4402	6	E30219
20	1016.6	78.1	5417	6	AX538619
21	904.8	69.5	13575	5	GGDYS
22	886	68.1	4402	6	E30220
23	750.4	57.7	4075	6	E30221
24	746.2	57.4	3747	6	E30218
25	722.4	55.5	2654	5	FSCDYSTRO
26	692.6	53.2	3163	6	E30223
27	608.2	46.7	3521	5	AF339031
28	584	44.9	2110	9	HUMDMDDX
29	584	44.9	4658	9	BC028720
30	570.4	43.8	10705	10	RNAJ2967
31	567.2	43.6	3161	10	MMGTTRPH
32	567.2	43.6	11096	6	AX538584
33	567.2	43.6	11096	10	MMV12229
34	553.8	42.6	6045	6	A63605
35	553.8	42.6	6045	6	AR281528
36	553.8	42.6	6059	6	AX107972
37	553.8	42.6	10302	6	AX538583
38	553.8	42.6	10302	9	HSMDPS
39	553.8	42.6	10320	6	A63607
40	553.8	42.6	10320	6	AR281529
41	551	42.4	3499	9	HSU43519
42	551	42.4	5106	6	AX552248
43	546.8	42.0	3073	10	AF195788
44	546.8	42.0	3140	10	AF195787
45	529.8	40.7	1240	10	RNAPDY3

ALIGNMENTS

RESULT 1					
AX538620					
LOCUS	AX538620	5339 bp	DNA	linear	PAT 23-NOV-2002
DEFINITION	Sequence 40 from Patent WO0229056.				
ACCESSION	AX538620				
VERSION	AX538620.1	GI:25271166			
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1				
AUTHORS	Chamberlain,J.S. and Harper,S.Q.				
TITLE	Mini-dystrophin nucleic acid and peptide sequences				
JOURNAL	Patent: WO 0229056-A 40 11-APR-2002;				
	THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)				

FEATURES		Location/Qualifiers						
source		1..5339						
		/organism="synthetic construct"						
		/mol_type="genomic DNA"						
		/db_xref="taxon:32630"						
BASE COUNT		1638 a	1191 c	1187 g	1323 t			
ORIGIN								
Query Match								
Best Local Similarity 100.0%; Score 1301; DB 6; Length 5339;								
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;								
QY	1	CGACTTTCCAGCAGTTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA	60					
DB	1993	CGACTTTCCAGCAGTTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA	2052					
QY	61	AACTAAAGAACCTGTAATCATGACTACTCTTGAGACTGTACGATATTTCTGACAGACA	120					
DB	2053	AACTAAAGAACCTGTAATCATGACTACTCTTGAGACTGTACGATATTTCTGACAGACA	2112					
QY	121	GCCTTTGGAAGACTAGAGAACTCTACAGAGGCCAGAGAGTGCCTCTCTGAGGAGAG	180					
DB	2113	GCCTTTGGAAGACTAGAGAACTCTACAGAGGCCAGAGAGTGCCTCTCTGAGGAGAG	2172					
QY	181	AGCCAGAAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGA	240					
DB	2173	AGCCAGAAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGA	2232					
QY	241	AAAAATGAACTGCACTCGCTGCTGAGGAGAGAAATAGATGAGACCTTTGAAAGACT	300					
DB	2233	AAAAATGAACTGCACTCGCTGCTGAGGAGAGAAATAGATGAGACCTTTGAAAGACT	2292					
QY	301	CCAGGAATCTTAAGAGGCCACGGATGAGTGAACCTCAAGCTGCGCCAAAGCTGAGGTGAT	360					
DB	2293	CCAGGAATCTTAAGAGGCCACGGATGAGTGAACCTCAAGCTGCGCCAAAGCTGAGGTGAT	2352					
QY	361	CAAGGATCTCGGAGCCGCTGGGAGATCTCTCATTCAGTCTCTCCAGATCACTCTCGA	420					
DB	2353	CAAGGATCTCGGAGCCGCTGGGAGATCTCTCATTCAGTCTCTCCAGATCACTCTCGA	2412					
QY	421	GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCAA	480					
DB	2413	GAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTGAGCCACGTCAA	2472					
QY	481	TGACCTTGCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCCTGATAACCTCAGCAC	540					
DB	2473	TGACCTTGCTGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCCTGATAACCTCAGCAC	2532					
QY	541	TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGGTGCGCGTCGAGGACCGAGTCAG	600					
DB	2533	TCTGGAAGACCTGAACACCCAGATGGAAGCTTCTGAGGTGCGCGTCGAGGACCGAGTCAG	2592					
QY	601	GCAGCTGATGAAGCCACAGGACATTTGGTCCAGCATCTCAGACATTTCTTCCAGCTC	660					
DB	2593	GCAGCTGATGAAGCCACAGGACATTTGGTCCAGCATCTCAGACATTTCTTCCAGCTC	2652					
QY	661	TGTCAGGCTCCCTGGGAGAGGACATCTCGGCCAAACAAAGTGCCCTACTATATCAACCA	720					
DB	2653	TGTCAGGCTCCCTGGGAGAGGACATCTCGGCCAAACAAAGTGCCCTACTATATCAACCA	2712					
QY	721	CGAGACTCAAAACAATCTGTGGGACATCCCAAAATGACAGAGTCTTACAGTCTTTAGC	780					
DB	2713	CGAGACTCAAAACAATCTGTGGGACATCCCAAAATGACAGAGTCTTACAGTCTTTAGC	2772					
QY	781	TGACCTGTAATATGTCAAGATCTCAGCTTATAGGACTGCGCATGAACCTCCGAAGACTGCA	840					
DB	2773	TGACCTGTAATATGTCAAGATCTCAGCTTATAGGACTGCGCATGAACCTCCGAAGACTGCA	2832					
QY	841	GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCACTGCTGATGCTGAGTGGGACCA	900					
DB	2833	GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCACTGCTGATGCTGAGTGGGACCA	2892					

QY	901	CAACCTCAAGCAAAATGACAGCCCATCGATATCCTCCAGATTATTAATTTGTTGACCAC	960
DB	2893	CAACCTCAAGCAAAATGACAGCCCATCGATATCCTCCAGATTATTAATTTGTTGACCAC	2953
QY	961	TATTTATACCCGCTCGAGCAAGAGCACAAATTTGGTCAAAGTCCCTCTCTCTCGGTGGA	1020
DB	2953	TATTTATACCCGCTCGAGCAAGAGCACAAATTTGGTCAAAGTCCCTCTCTCTCGGTGGA	3012
QY	1021	TATGTGCTCGAACTGGCTGCTGAATGTTTATGATACGGACGACAGGAGGATCCGTGT	1080
DB	3013	TATGTGCTCGAACTGGCTGCTGAATGTTTATGATACGGACGACAGGAGGATCCGTGT	3072
QY	1081	CCTGTCCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAACACAAGTACAG	1140
DB	3073	CCTGTCCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAACACAAGTACAG	3132
QY	1141	ATACCTTTTCAAGCAAGTGGCAAGTTCAAACAGATTTTGTGACAGCGCAGGCTGGGCCT	1200
DB	3133	ATACCTTTTCAAGCAAGTGGCAAGTTCAAACAGATTTTGTGACAGCGCAGGCTGGGCCT	3192
QY	1201	CCTTCTGCATGATCTTCTTCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCTTTGGGGG	1260
DB	3193	CCTTCTGCATGATCTTCTTCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCATCTTTGGGGG	3252
QY	1261	CAGTAACATTGAGCCAAAGTTCGGAGCTGCTTCCCAATTTG	1301
DB	3253	CAGTAACATTGAGCCAAAGTTCGGAGCTGCTTCCCAATTTG	3293

RESULT 2

AX538621

LOCUS

Sequence 41 from Patent WO229056.

AX538621

ACCESSION

AX538621.1

VERSION

GI:25271168

KEYWORDS

synthetic construct

SOURCE

synthetic construct

ORGANISM

artificial sequences.

REFERENCE

1

AUTHORS

Chamberlain, J.S. and Harper, S.Q.

TITLE

Mini-dystrophin nucleic acid and peptide sequences

JOURNAL

Patent: WO 0229056-A 41 11-APR-2002;

THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES

Location/Qualifiers

1..5462

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="Synthetic"

BASE COUNT

1668 a 1225 c 1212 g 1357 t

ORIGIN

Query Match

100.0%; Score 1301; DB 6; Length 5462;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1

CGACTTTCCAGCAGTTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA

60

DB

2116

CGACTTTCCAGCAGTTTCAGAACGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA

2175

QY

61

AACTAAAGAACCTGTAATCATGACTACTCTTGAGACTGTACGATATTTCTGACAGACA

120

DB

2176

AACTAAAGAACCTGTAATCATGACTACTCTTGAGACTGTACGATATTTCTGACAGACA

2235

QY

121

GCCTTTGGAAGACTAGAGAACTCTTACAGAGGCCAGAGAGTGCCTCTCTGAGGAGAG

180

DB

2236

GCCTTTGGAAGACTAGAGAACTCTTACAGAGGCCAGAGAGTGCCTCTCTGAGGAGAG

2295

QY

181

AGCCAGAAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGA

240

DB

2296

AGCCAGAAATGTCACTCGGCTTTCTACGAAAGCAGGCTGAGAGGTCATATCTGAGTGGGA

2355

241 AAAATTGAACCTGACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACT 300
2356 AAAATTGAACCTGACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAAGACT 2415
301 CCAGGAACCTCAAGAGCCACCGATAGCTGAGACTCAAGCTGCGCCCAAGCTGAGGTGAT 360
2416 CCAGGAACCTCAAGAGCCACCGATAGCTGAGACTCAAGCTGCGCCCAAGCTGAGGTGAT 2475
361 CAAGGGATCTGGCAGCCGCTGGGCGGATCTCTCATTTGATCTCTCCAGATCACTCTCGA 420
2476 CAAGGGATCTGGCAGCCGCTGGGCGGATCTCTCATTTGATCTCTCCAGATCACTCTCGA 2535
421 GAAAGTCAAGGACCTTCAGAGAGAAATGGCCCTCTCAAGAGAACTGAGAGCAAGTCAA 480
2536 GAAAGTCAAGGACCTTCAGAGAGAAATGGCCCTCTCAAGAGAACTGAGAGCAAGTCAA 2595
481 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCCAGTATCACTCAGCAC 540
2596 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCCAGTATCACTCAGCAC 2655
541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCCTCGAGGACCGAGTCA 600
2656 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCCTCGAGGACCGAGTCA 2715
601 GCAGCTGATGAAGCCACACAGGACCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTC 660
2716 GCAGCTGATGAAGCCACACAGGACCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTC 2775
661 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 720
2776 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 2835
721 CGAGACTCAACAACTTGTCTGGGACCAATCCCAAAATGACAGAGCTTACAGCTCTTTAGC 780
2836 CGAGACTCAACAACTTGTCTGGGACCAATCCCAAAATGACAGAGCTTACAGCTCTTTAGC 2895
781 TGACCTGAATAATGTGAGATCTCAGCTTATAGAGCTGCCATGAATCTCCGAGAGCTGCA 840
2896 TGACCTGAATAATGTGAGATCTCAGCTTATAGAGCTGCCATGAATCTCCGAGAGCTGCA 2955
841 GAAAGCCCTTTGCTTGGATCTTTGAGCTGTGAGCTGATGATGATGATGATGATGATGAT 900
2956 GAAAGCCCTTTGCTTGGATCTTTGAGCTGTGAGCTGATGATGATGATGATGATGATGAT 3015
901 CAACCTCAAGCAAAATGACAGCCATGATATCTCAGATTAATTAATTTGTTGACCA 960
3016 CAACCTCAAGCAAAATGACAGCCATGATATCTCAGATTAATTAATTTGTTGACCA 3075
961 TAATTTATGACCGCTGGAGCAAGAGCAACAATTTGTTGATCAAGCTCCCTCTCTGCGTGA 1020
3076 TAATTTATGACCGCTGGAGCAAGAGCAACAATTTGTTGATCAAGCTCCCTCTCTGCGTGA 3135
1021 TATGTGTCTGAACCTGGCTGTGAATGTTTATGATGATGATGATGATGATGATGATGATGAT 1080
3136 TATGTGTCTGAACCTGGCTGTGAATGTTTATGATGATGATGATGATGATGATGATGATGAT 3195
1081 CCTGTCTTTTAAAGTGCATCAATTTCCCTGTGTAAGAGCAATTTGGAAGACAGTACAG 1140
3196 CCTGTCTTTTAAAGTGCATCAATTTCCCTGTGTAAGAGCAATTTGGAAGACAGTACAG 3255
1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 1200
3256 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 3315
1201 CCTTCTGATGATCTATCAAAATTCAGAGAGTGGGTGAGTTCATCTTTTGGGG 1260
3316 CCTTCTGATGATCTATCAAAATTCAGAGAGTGGGTGAGTTCATCTTTTGGGG 3375
1261 CAGTAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
3376 CAGTAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTTG 3416

RESULT 3
AR304538
LOCUS AR304538
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 5952)
Xiao, X. and Liu, P. X.
METHOD Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES
Location/Qualifiers
1..5952
/organism="unknown"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 5952;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGCAGCTTCAGAGAGCAACGATGTACATAGGCGCTTCAGAGGGGATTCGA 60
DB 3402 CGACTTTCAGCAGCTTCAGAGAGCAACGATGTACATAGGCGCTTCAGAGGGGATTCGA 3461
QY 61 AACTTAAAGAACCTGTAACTCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACA 120
DB 3462 AACTTAAAGAACCTGTAACTCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACA 3521
QY 121 GCCTTTGAGAGACTAGAGAAATCTTACAGAGCCACAGAGAGTGCCTCTGAGGAGAG 180
DB 3522 GCCTTTGAGAGACTAGAGAAATCTTACAGAGCCACAGAGAGTGCCTCTGAGGAGAG 3581
QY 181 AGCCAGAGATGTCACCTCGCTTCTACGAAAGCAGGCTGAGAGGTCAATATCTGAGTGGGA 240
DB 3582 AGCCAGAGATGTCACCTCGCTTCTACGAAAGCAGGCTGAGAGGTCAATATCTGAGTGGGA 3641
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 300
DB 3642 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 3701
QY 301 CCAGGAACCTCAAGAGCCACCGATGAGCTGAGCTCAAGCTGCGCCAGCTGAGGTGAT 360
DB 3702 CCAGGAACCTCAAGAGCCACCGATGAGCTGAGCTCAAGCTGCGCCAGCTGAGGTGAT 3761
QY 361 CAAGGGATCTGGCAGCCGCTGGGCGGATCTCTCATTTGATCTCTCCAGATCACTCTCGA 420
DB 3762 CAAGGGATCTGGCAGCCGCTGGGCGGATCTCTCATTTGATCTCTCCAGATCACTCTCGA 3821
QY 421 GAAAGTCAAGGACCTTCAGAGAGAAATTTGGCCCTCTGAAAGAGAACGTGAGCCAGTCAA 480
DB 3822 GAAAGTCAAGGACCTTCAGAGAGAAATTTGGCCCTCTGAAAGAGAACGTGAGCCAGTCAA 3881
QY 481 TGACCTTGCTGCGCAGCTTACCACTTTTGGGCAATTCAGCTCTCACGATATAACCTCAGCAC 540
DB 3882 TGACCTTGCTGCGCAGCTTACCACTTTTGGGCAATTCAGCTCTCACGATATAACCTCAGCAC 3941
QY 541 TCTGGAAGACCTGGAACACCAAGATGGAAGCTTCTGAGGTGGCCCTCGAGGACCGAGTCA 600
DB 3942 TCTGGAAGACCTGGAACACCAAGATGGAAGCTTCTGAGGTGGCCCTCGAGGACCGAGTCA 4001
QY 601 GCAGCTGATGAAGCCACAGGACTTTGGTCCAGATCTCAGCACTCTCAGCACTTTCTTTCCAGCTC 660
DB 4002 GCAGCTGATGAAGCCACAGGACTTTGGTCCAGATCTCAGCACTCTCAGCACTTTCTTTCCAGCTC 4061
QY 661 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 720
DB 4062 TGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACCA 4121

QY	721	CGAGACTCAAACTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC	780
Db	4122	CGAGACTCAAACTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC	4181
QY	781	TGACCTGGAATAATGTCAGATTTCTAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA	840
Db	4182	TGACCTGGAATAATGTCAGATTTCTAGCTTATAGGACTGCCATGAACTCCGAAGACTGCA	4241
QY	841	GAAGGCCCTTTGCTGGATCTCTTGGAGCTGTCAGCTGATGATGCTTTGGACACCA	900
Db	4242	GAAGGCCCTTTGCTGGATCTCTTGGAGCTGTCAGCTGATGATGCTTTGGACACCA	4301
QY	901	CAACCTCAAGCAAAATGACAGGCCATGATATCTGACAGATTAATTAATGTTTGACCAAC	960
Db	4302	CAACCTCAAGCAAAATGACAGGCCATGATATCTGACAGATTAATTAATGTTTGACCAAC	4361
QY	961	TATTTATGACCGCTGGAGCAAGACCAAAATTTGGTCAAGCTCTCTCTGGTGGG	1020
Db	4362	TATTTATGACCGCTGGAGCAAGACCAAAATTTGGTCAAGCTCTCTCTGGTGGG	4421
QY	1021	TATGCTGCTGAACCTGGCTGCTGAATGTTTATGATACGGACGAAACAGGAGGATCCGTGT	1080
Db	4422	TATGCTGCTGAACCTGGCTGCTGAATGTTTATGATACGGACGAAACAGGAGGATCCGTGT	4481
QY	1081	CTCTGCTTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG	1140
Db	4482	CTCTGCTTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG	4541
QY	1141	ATACCTTTTCAAGCAAGTGCAGATTTCAACAGGATTTTGTACCGACGAGGCTGGGCT	1200
Db	4542	ATACCTTTTCAAGCAAGTGCAGATTTTGTACCGACGAGGCTGGGCT	4601
QY	1201	CTTCTGATGATTTTATCCAAATTTCCAGACAGTTGGGTGAAGTTGCATCTTTGGGG	1260
Db	4602	CTTCTGATGATTTTATCCAAATTTCCAGACAGTTGGGTGAAGTTGCATCTTTGGGG	4661
QY	1261	CAGTAACATTGAGCAAGTGTCCGGAGCTCTTCCCAATTTG	1301
Db	4662	CAGTAACATTGAGCAAGTGTCCGGAGCTCTTCCCAATTTG	4702
RESULT 4			
LOCUS	AX114289	5952 bp	DNA linear PAT 11-MAY-2001
DEFINITION	Sequence 1 from Patent WO0129243.		
ACCESSION	AX114289		
VERSION	AX114289.1	GI:14031259	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Paul, X. L. and Xiao, X.		
TITLE	Method and vector for producing and transferring trans -spliced peptides		
JOURNAL	Patent: WO 0129243-A 1 26-APR-2001; DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION (US)		
FEATURES	Location/Qualifiers		
source	1..5952		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
misc_feature	2897..2898		
	/note="S4 junction site"		
misc_feature	3198..3199		
	/note="S2 junction site"		
BASE COUNT	1860 a 1344 c 1410 g 1338 t		
ORIGIN			
Query Match	100.0%; Score 1301; DB 6; Length 5952;		
Best Local Similarity	100.0%; Pred. No. 0;		

Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CGACTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA	60
Db	3402	CGACTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA	3461
QY	61	AACCTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGCA	120
Db	3462	AACCTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAAATATTTCTGACAGCA	3521
QY	121	GCCTTTGGAAGACTAGAAACTCTTACAGAGCCAGAGAGCTCCCTCTGAGAGAG	180
Db	3522	GCCTTTGGAAGACTAGAAACTCTTACAGAGCCAGAGAGCTCCCTCTGAGAGAG	3581
QY	181	AGCCCAAGATGTCACTCGGCTTCTACAAAGCAGGCTGAGGAGGTCAATCTAGTGGGA	240
Db	3582	AGCCCAAGATGTCACTCGGCTTCTACAAAGCAGGCTGAGGAGGTCAATCTAGTGGGA	3641
QY	241	AAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT	300
Db	3642	AAATTTGAACCTGCACCTCCGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT	3701
QY	301	CCAGGAATTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGTGT	360
Db	3702	CCAGGAATTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGTGT	3761
QY	361	CAAGGATTCCTGGAGCCCTCGCTGACTGGCAGAGATCTCTCATGACTCTCTCCAGATCACCTCGA	420
Db	3762	CAAGGATTCCTGGAGCCCTCGCTGACTGGCAGAGATCTCTCATGACTCTCTCCAGATCACCTCGA	3821
QY	421	GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCTCTGAAAGAGAACTGAGCCACCTCAA	480
Db	3822	GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCTCTGAAAGAGAACTGAGCCACCTCAA	3881
QY	481	TGACCTTCTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCCCGTATTAACCTCAGCAC	540
Db	3882	TGACCTTCTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCCCGTATTAACCTCAGCAC	3941
QY	541	TCTCGAAGACTGGAACACCAAGTGGAGCTTCTGAGGTGGCGTGGAGCCGAGCTCAG	600
Db	3942	TCTCGAAGACTGGAACACCAAGTGGAGCTTCTGAGGTGGCGTGGAGCCGAGCTCAG	4001
QY	601	GCAGCTGATGAAGCCCAAGGACTTTTGGTCAGCATCTCAGACCTTTCTTCCAGCTC	660
Db	4002	GCAGCTGATGAAGCCCAAGGACTTTTGGTCAGCATCTCAGACCTTTCTTCCAGCTC	4061
QY	661	TGTCAGGCTCCCTGGGAGAGCCATCTCGCAGAAACAAAGTGCCCTACTATATCAACCA	720
Db	4062	TGTCAGGCTCCCTGGGAGAGCCATCTCGCAGAAACAAAGTGCCCTACTATATCAACCA	4121
QY	721	CGAGACTCAAACTTGTCTGGGACCATCCCAAAATGACAGAGCTTACAGCTTTTAGC	780
Db	4122	CGAGACTCAAACTTGTCTGGGACCATCCCAAAATGACAGAGCTTACAGCTTTTAGC	4181
QY	781	TGACCTGAAATATGACAGTTCTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCA	840
Db	4182	TGACCTGAAATATGACAGTTCTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCA	4241
QY	841	GAAGGCCCTTTGGTGGATCTCTTGGAGCTGTCAGCTGATGATGCTTTGGACACCA	900
Db	4242	GAAGGCCCTTTGGTGGATCTCTTGGAGCTGTCAGCTGATGATGCTTTGGACACCA	4301
QY	901	CAACCTCAAGCAAAATGACAGGCCATGATATCTGACAGATTAATTAATGTTTGACAC	960
Db	4302	CAACCTCAAGCAAAATGACAGGCCATGATATCTGACAGATTAATTAATGTTTGACAC	4361
QY	961	TATTTATGACCGCTGGAGCAAGACCAAAATTTGGTCAACGTCCCTCTCTGGTGGG	1020
Db	4362	TATTTATGACCGCTGGAGCAAGACCAAAATTTGGTCAACGTCCCTCTCTGGTGGG	4421
QY	1021	TATGCTGCTGAACCTGGCTGCTGAATGTTATGATACGGGACGAAACAGGGAGGATCCGTGT	1080
Db	4422	TATGCTGCTGAACCTGGCTGCTGAATGTTATGATACGGGACGAAACAGGGAGGATCCGTGT	4481

QY 1081 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 1140
DB 4482 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 4541
QY 1141 ATACCTTTTCAAGCAAGTGCAGAGTTCACAGAGATTTGTGACAGCGCAGGCTGGGCT 1200
DB 4542 ATACCTTTTCAAGCAAGTGCAGAGTTCACAGAGATTTGTGACAGCGCAGGCTGGGCT 4601
QY 1201 CCTTCTGCATGATTCATCCAAATTCACAGACAGTGGGTGAAGTTGCATCCTTTGGGG 1260
DB 4602 CCTTCTGCATGATTCATCCAAATTCACAGACAGTGGGTGAAGTTGCATCCTTTGGGG 4661
QY 1261 CAGTAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
DB 4662 CAGTAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTTG 4702

RESULT 5
AX538622
LOCUS
DEFINITION
Sequence 42 from Patent WO0229056.
ACCESSION
AX538622
VERSION
AX538622.1 GI:25271171
KEYWORDS
synthetic construct
ORGANISM
synthetic construct
artificial sequences.

REFERENCE
1
AUTHORS
Chamberlain, J.S. and Harper, S.Q.
TITLE
Mini-dystrophin nucleic acid and peptide sequences
JOURNAL
Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..8689
/organism="synthetic construct"
/mol type="genomic DNA"
/db xref="taxon:32630"
/note="Synthetic"

BASE COUNT 2721 a 1804 c 1861 g 2303 t
ORIGIN

Query Match 100.0%; Score 1301; DB 6; Length 8689;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
DB 3448 CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 3507

QY 61 AACTTAAAGAACCTGTATCATGTACTCTTTGAGACTGTACGAATATTTCTGACAGACGA 120
DB 3508 AACTTAAAGAACCTGTATCATGTACTCTTTGAGACTGTACGAATATTTCTGACAGACGA 3567

QY 121 GCCTTTTGGAGGACTAGAGAACTCTACAGGAGCCGACAGAGCTGCTCCTGAGGAGAG 180
DB 3568 GCCTTTTGGAGGACTAGAGAACTCTACAGGAGCCGACAGAGCTGCTCCTGAGGAGAG 3627

QY 181 AGCCAGAAATGTCACTCGGCTTCTACAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA 240
DB 3628 AGCCAGAAATGTCACTCGGCTTCTACAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA 3687

QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGAAGCT 300
DB 3688 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGAAGCT 3747

QY 301 CCAGGAACCTTCAAGAGCCGCGATGAGCTGAGCTCAAGCTGCGGCGACAGCTGAGTGTAT 360
DB 3748 CCAGGAACCTTCAAGAGCCGCGATGAGCTGAGCTCAAGCTGCGGCGACAGCTGAGTGTAT 3807

QY 361 CAAGGGATCCTGGCAGCCGCTGGGCGGATCTCTCATTTGACTCTCTCAAGATCACTCCGA 420
DB 3808 CAAGGGATCCTGGCAGCCGCTGGGCGGATCTCTCATTTGACTCTCTCAAGATCACTCCGA 4867

QY 421 GAAAGTCAAGSCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 480
DB 3868 GAAAGTCAAGSCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACGTTGAGCCACGTCAA 3927
QY 481 TGACCTTCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCAACGTATTAACCTCAGCAC 540
DB 3928 TGACCTTCTCGCCAGCTTACCACTTTTGGGCAATTCAGCTCTCAACGTATTAACCTCAGCAC 3987
QY 541 TCTGGAAGACCTGGAACACCCAGATGGAAGCTTCTGAGGTGCGCGTCGAGGACCGAGTCAG 600
DB 3988 TCTGGAAGACCTGGAACACCCAGATGGAAGCTTCTGAGGTGCGCGTCGAGGACCGAGTCAG 4047
QY 601 GCAGCTGCATGAAGCCCAACAGGAGCTTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 660
DB 4048 GCAGCTGCATGAAGCCCAACAGGAGCTTTTGGTCCAGCATCTCAGCACCTTTCTTCCACGTC 4107
QY 661 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
DB 4108 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 4167
QY 721 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 780
DB 4168 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGC 4227
QY 781 TGACCTGAATTAATGTGAGATTTCTGAGCTTATAGACTGCGCATGAAACTCCGAGACTGCA 840
DB 4228 TGACCTGAATTAATGTGAGATTTCTGAGCTTATAGACTGCGCATGAAACTCCGAGACTGCA 4287
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGTGATGCTTGGACCAAGCA 900
DB 4288 GAAGGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGTGATGCTTGGACCAAGCA 4347
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTAATTAATTTGTTGACCA 960
DB 4348 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTAATTAATTTGTTGACCA 4407
QY 961 TATTTATGACCGCTGGGAGCAAGAGCACAAATTTGCTCAACGCTCCCTCTCTGCGTGGGA 1020
DB 4408 TATTTATGACCGCTGGGAGCAAGAGCACAAATTTGCTCAACGCTCCCTCTCTGCGTGGGA 4467
QY 1021 TATGCTCTGAATCGCTGCTGTAATGTTTATGATACGGGACGAACAGGGAGGATCCCTGT 1080
DB 4468 TATGCTCTGAATCGCTGCTGTAATGTTTATGATACGGGACGAACAGGGAGGATCCCTGT 4527
QY 1081 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 1140
DB 4528 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAG 4587
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGGCTGGGCT 1200
DB 4588 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGGCTGGGCT 4647
QY 1201 CCTTCTGCATGATTCATCCAAATTCACAGACAGTGGGTGAAGTTGCATCCTTTGGGG 1260
DB 4648 CCTTCTGCATGATTCATCCAAATTCACAGACAGTGGGTGAAGTTGCATCCTTTGGGG 4707
QY 1261 CAGTAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
DB 4708 CAGTAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTTG 4748

RESULT 6
AX538624
LOCUS
DEFINITION
Sequence 44 from Patent WO0229056.
ACCESSION
AX538624
VERSION
AX538624.1 GI:25271175
KEYWORDS
synthetic construct
ORGANISM
synthetic construct
artificial sequences.

REFERENCE
1

AUTHORS Chamberlain, J.S. and Harper, S.O.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 44 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1. .11443
/organism="synthetic construct"
/mol type="genomic DNA"
/db xref="taxon:32630"
/notes="Synthetic"
BASE COUNT 3707 a 2339 c 2502 g 2895 t
ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 11443;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGAGCTTCAGAGCAGAACGATGTACATAGGGCTTCAGAGGGAAATTGAA 60
DB 6202 CGACTTTCAGAGCTTCAGAGCAGAACGATGTACATAGGGCTTCAGAGGGAAATTGAA 6261
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTCAGACAGCA 120
DB 6262 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTCAGACAGCA 6321
QY 121 GCCTTTGGAGAGCTAGAGAACTCTACAGAGGCCAGAGAGCTGCCCTCTCGAGGAGAG 180
DB 6322 GCCTTTGGAGAGCTAGAGAACTCTACAGAGGCCAGAGAGCTGCCCTCTCGAGGAGAG 6381
QY 181 AGCCAGAGTGCACCTCGCTTCTAGAAAGCAGAGCTGAGGAGGTCATPACTGAGTGGGA 240
DB 6382 AGCCAGAGTGCACCTCGCTTCTAGAAAGCAGAGCTGAGGAGGTCATPACTGAGTGGGA 6441
QY 241 AAAATTGAACCTGCACCTCGCTGACCTGGCAGAGAAAATAGATGAGACCTTTGAAGACT 300
DB 6442 AAAATTGAACCTGCACCTCGCTGACCTGGCAGAGAAAATAGATGAGACCTTTGAAGACT 6501
QY 301 CCAGGAATTCAGAGGCCAGATGAGCTGAGCTTCAAGCTTCCGCAAGCTTGAGGTGAT 360
DB 6502 CCAGGAATTCAGAGGCCAGATGAGCTGAGCTTCAAGCTTCCGCAAGCTTGAGGTGAT 6561
QY 361 CAAGGATCTCGCAGCCCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 420
DB 6562 CAAGGATCTCGCAGCCCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 6621
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCTTCTGAAAGAGAGCTGAGCCAGCTGAA 480
DB 6622 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCTTCTGAAAGAGAGCTGAGCCAGCTGAA 6681
QY 481 TGACCTTGTCCGACGCTTACCACTTTGGCATTTCAGCTCTCACCGTATAAAGCTCAGAC 540
DB 6682 TGACCTTGTCCGACGCTTACCACTTTGGCATTTCAGCTCTCACCGTATAAAGCTCAGAC 6741
QY 541 TCTGGAAGACTGGAACACAGATGGAAGCTTTCGAGGTGGCGCTCGAGGACCGAGTCCAG 600
DB 6742 TCTGGAAGACTGGAACACAGATGGAAGCTTTCGAGGTGGCGCTCGAGGACCGAGTCCAG 6801
QY 501 GCAGCTGCATGAAGCCACAGGAGCTTGTTCAGGATCTCAGCATCTTCAGCTTCTTTCACGCT 660
DB 6802 GCAGCTGCATGAAGCCACAGGAGCTTGTTCAGGATCTCAGCATCTTCAGCTTCTTTCACGCT 6861
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720
DB 6862 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 6921
QY 721 CGAGACTCAACAACTTGTGGGAGCCATCCCAAAATGACAGAGCTTACAGCTCTTTAGC 780
DB 6922 CGAGACTCAACAACTTGTGGGAGCCATCCCAAAATGACAGAGCTTACAGCTCTTTAGC 6981
QY 781 TGACCTGAATATCTCAGATTTCTCAGCTTATAGGATGTCATGAAACTCCGAGACTGCA 840
DB 6982 TGACCTGAATATCTCAGATTTCTCAGCTTATAGGATGTCATGAAACTCCGAGACTGCA 7041

841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACAGCA 900
DB 7042 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCTTGGACAGCA 7101
QY 901 CAACCTCAGCAAAATGACAGCCCATGATATCTTGGATATCTTGGATATTAATTTGTTGACCA 960
DB 7102 CAACCTCAGCAAAATGACAGCCCATGATATCTTGGATATCTTGGATATTAATTTGTTGACCA 7161
QY 961 TATTATGACCGCTCGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 1020
DB 7162 TATTATGACCGCTCGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGA 7221
QY 1021 TATGTGTCGAATGCTGCTGAATGTTTATGATACGGAGCAAGAGGATCGGTGT 1080
DB 7222 TATGTGTCGAATGCTGCTGAATGTTTATGATACGGAGCAAGAGGATCGGTGT 7281
QY 1081 CCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAAGCACATTTTGGAAAGACAAGTACAG 1140
DB 7282 CCTGTCTTTTAAACTGGCATCATTTTCCCTGTGTAAGCACATTTTGGAAAGACAAGTACAG 7341
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCCAGCGAGGCTGGGCT 1200
DB 7342 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCCAGCGAGGCTGGGCT 7401
QY 1201 CCTTCTGCATGATCTTCCAAATTTCCAGACAGTTGGGTGAAGTTGCATCTTTTGGGG 1260
DB 7402 CCTTCTGCATGATCTTCCAAATTTCCAGACAGTTGGGTGAAGTTGCATCTTTTGGGG 7461
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 1301
DB 7462 CAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTG 7502

RESULT 7
AX538627 12057 bp DNA linear PAT 23-NOV-2002
LOCUS Sequence 47 from Patent WO0229056.
DEFINITION
AX538627
VERSION AX538627.1 GI:25271181
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.O.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1. .12057
/organism="synthetic construct"
/mol type="genomic DNA"
/db xref="taxon:32630"
/notes="Synthetic"
BASE COUNT 4020 a 2456 c 2790 g 2791 t
ORIGIN
Query Match 100.0%; Score 1301; DB 6; Length 12057;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGAGCTTCAGAGCAGAACGATGTACATAGGGCTTCAGAGGGAAATTGAA 60
DB 8716 CGACTTTCAGAGCTTCAGAGCAGAACGATGTACATAGGGCTTCAGAGGGAAATTGAA 8775
QY 61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 120
DB 8776 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCA 8835
QY 121 GCCTTTGGAGAGCTAGAGAACTCTTACAGGAGGCCAGAGAGCTGCCCTCTCGAGGAGAG 180
DB 8836 GCCTTTGGAGAGCTAGAGAACTCTTACAGGAGGCCAGAGAGCTGCCCTCTCGAGGAGAG 8895

181 AGCCAGAAATGTCACCTGGCTTCTCAAGAAAGCAGGCTGAGGAGGTCAATACCTAGTGGGA 240
8896 AGCCAGAAATGTCACCTGGCTTCTCAAGAAAGCAGGCTGAGGAGGTCAATACCTAGTGGGA 8955
241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAGACT 300
8956 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAGACT 9015
301 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGTGAT 360
9016 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGTGAT 9075
361 CAAGGGATCTCGGAGCCCTGGGCGATCTCTCATATGACTCTCTCAAGATCACTCGA 420
9076 CAAGGGATCTCGGAGCCCTGGGCGATCTCTCATATGACTCTCTCAAGATCACTCGA 9135
421 GAAAGTCAAGGCACTTCGAGAGAAATGCGCTCTGAAGAGAAAGCTGAGCAGCACTCAA 480
9136 GAAAGTCAAGGCACTTCGAGAGAAATGCGCTCTGAAGAGAAAGCTGAGCAGCACTCAA 9195
481 TGACCTTGCTGGCAGCTTACCACTTTGGGCACTTCAAGCTTCAAGCTTCAAGCTTCAAG 540
9196 TGACCTTGCTGGCAGCTTACCACTTTGGGCACTTCAAGCTTCAAGCTTCAAGCTTCAAG 9255
541 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGAGCGAGTCAAG 600
9256 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGAGCGAGTCAAG 9315
601 GCAGCTGCATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
9316 GCAGCTGCATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 9375
661 TGTCAGGCTTCTGGGAGAGCCATCTCGCAACAAAGTGGCTTCTATATCAACCA 720
9376 TGTCAGGCTTCTGGGAGAGCCATCTCGCAACAAAGTGGCTTCTATATCAACCA 9435
721 CGAGACTCAAAACAACTTCTGGGAGAGCCATCTCGCAACAAAGTGGCTTCTATATCAACCA 780
9436 CGAGACTCAAAACAACTTCTGGGAGAGCCATCTCGCAACAAAGTGGCTTCTATATCAACCA 9495
781 TGACCTGAATATGTCAGATCTCAGCTTATAGAGCTGCGCATGAATCTCGAAGACTGCA 840
9496 TGACCTGAATATGTCAGATCTCAGCTTATAGAGCTGCGCATGAATCTCGAAGACTGCA 9555
841 GAAGGCCCTTGTGCTGATCTTGTAGCCTTGTGATGATGATGATGATGATGATGATGATGAT 900
9556 GAAGGCCCTTGTGCTGATCTTGTAGCCTTGTGATGATGATGATGATGATGATGATGATGAT 9615
901 CAACCTCAAGCAAAATGACCAAGCCATGATATCTCTGAGATTTATTAATTTGTTGACCA 960
9616 CAACCTCAAGCAAAATGACCAAGCCATGATATCTCTGAGATTTATTAATTTGTTGACCA 9675
961 TATTTATGAGCGCTGAGCAAGAGCAACAAATTTGGTCAAGCTCTCTGCGTGA 1020
9676 TATTTATGAGCGCTGAGCAAGAGCAACAAATTTGGTCAAGCTCTCTGCGTGA 9735
1021 TATGTGCTGAACCTGCTGATGATTTATGATGATGATGATGATGATGATGATGATGATGAT 1080
9736 TATGTGCTGAACCTGCTGATGATTTATGATGATGATGATGATGATGATGATGATGATGAT 9795
1081 CCTGTCTTTAAATCTGGATCATTTCCCTGTTGTAAGACATTTGGAAGCAAGTACAG 1140
9796 CCTGTCTTTAAATCTGGATCATTTCCCTGTTGTAAGACATTTGGAAGCAAGTACAG 9855
1141 ATACCTTTTCAAGCAAGTGGCAAGTCTCAAGATTTTGTGACAGCGAGGCTGGGCT 1200
9856 ATACCTTTTCAAGCAAGTGGCAAGTCTCAAGATTTTGTGACAGCGAGGCTGGGCT 9915
1201 CCTTCTGCATGATTTCTATCAAAATTCAGACAGTGGGTGAAGTTCATCTTTGGGG 1260
9916 CCTTCTGCATGATTTCTATCAAAATTCAGACAGTGGGTGAAGTTCATCTTTGGGG 9975
1261 CAGTAAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTG 1301

Db 9976 CAGTAAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTG 10016
RESULT 8
HSDMR Human mRNA for dystrophin. 12446 bp mRNA linear PRI 12-SEP-1993
LOCUS X14298
DEFINITION GI:30845
ACCESSION X14298.1
VERSION Dmd gene; Duchenne muscular dystrophy; dystrophin.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
REFERENCE
AUTHORS Rosenthal, A., Speer, A., Billwitz, H., Cross, G.S., Forrest, S.M. and Davies, K.E.
TITLE Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous
JOURNAL Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE 89345106
PUBMED 2668885
REFERENCE 2 (bases 1 to 12446)
AUTHORS Rosenthal, A.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1989) Rosenthal, A., Akademie der Wissenschaften der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Rössle Str.10, 1115 Berlin Buch, DDR
COMMENT see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.
FEATURES
source
1..12446
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X chromosomal, Xp21."
/tissue_type="muscle"
/dev_stage="fetal and adult."
99..11156
/note="dystrophin (AA 1 - 3685)"
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/db_xref="SWISS-PROT:P11532"
/translation="MLWEEVEDEYEDYVQKKTFTKWNVAQSKGKHENLPSDL QGRRLLDLLEGUTGQKPKREKSTRVHALNNNNKALRVLQNNVDLVNIGSDIVDG NKLTLGLIWNILHWQVKNMKNIMAGLOFTNSEKILLSWVRQSTNRYPOVNVINFT TSWSDGLANALIHSHRPDLFDWNSVVCQSQATQRLHAFNIARYQLGIGKLDPEDY DTTVPDKSILMYITSLFQVLPOQVSTEALQEVEMLPKPKVKEHFQHLHQHVSQ QITVSLAQDYERTSSPKRPFKSVAYTQAAVTTSDPTRSPPOHFEAPDKHFGSSL MESEVNLDRYCTALEVLSLNAEDTLOAGREISNDVEVVKDQFHTHEGYMDLTAH CQRVGNILQIGSKLIGTGKLSDEETVEQOMNLLNSRWECILVASKEOSNLRHVM DLQNKULELNDLWTKTEETRKMEPEPLQDLEDKRQVQKHVLEDELEQQRVYN ECLFSAWLSKEDAVNKIHTGFDQNMELSLQKLVADKADLEKQSGMGLYSIK EQLLTLTKNSVQKTEAWLNDNFARCDNLVQKLEKSTQISQAVTTTQPSLTFTVM ETVTTTTRQILVKHAEELPPPPPOKQKITVDSIRKRLVDVITELHSWITRSEA VLOSPEAJFRKGNFSLDKENALEREKAKERKLODASRSGQALVEOMNEGVNA DSIQASEQLNSRIEFCQLLSERLWLYQNNILAFYNQLOLEQMTTAAENWLKIQ PTTSEPTAKSKQIKICQENRSLGQPQIERKIQSTALKKQSGPMELDADFVAF TNHFQKFSVDVQAREKELQITFDLPMPRYOETMSAIRTWQSETKLSIPQUSVTDY EIMEQRLQALQSSLEQEQSGLYLSTTKEMSKKAPSEIRKQSEFEELEGWK KCLSLVHEQKLEEQNMKNRKIQNHITLQKWAEDVDFKEEPALGSGEILKKQL KOCRLVSDICTQPSLNSVNGGOKIKNEAPEFASRLTELKELNTOWDHMCQVY EAKALQKAKVKLLTESVNSVIAQAPVQAQKLELTITNYQMLCTRLNGKKTLE EBAQCKWELLSYLEKANKWLNVEFKLITENIPGABEISEVLSLNLNMRSEDPN NQIRILAQITLTDGVMDELINSELTNSRWELHEAVRRQKLEESIOSAQGETNS LHLTQESLITIDKLAAYADKYDAAPQEQAKIOSDLTSHSILEEMCKHQGKEA AQRVLSQIDVQAQKLDQVMKFLFKPANKFEQKMLIDVEMKMLPALLETKSYE QEVVQSQNLNHCNLYKLSLVKSEVEMVINTGRQIVQKQKTENPKELDERLTALKHY

NELGAKVTERKQLEKICLILSRKMKEMVNLTEWLAATDMELTKRSAVEGMPNLSLE
VANGAKOKELEKQKHLKSTVEVEGALKTVLGGKLTVEDKLSLNSNWIATVSAE
EWNLLETRDQAAALNHNHGHCKRLVEPQISLELLEDESEKKPKQKEDYKRLKABELND
IRPKVDSTRDQAAALNHNHGHCKRLVEPQISLELLEDESEKKPKQKEDYKRLKABELND
QFNSDIQKLEPTEABTQOQVYNLEDFNKMEDNEGTVKLELQORDNLOQRIITELER
KSEBIIKQIQLLOTQKHALDLRQORAKALEI SHOWYKQRQADLLKLDITIEKL
ASLEPPREDRIKIDIRELQKKEELNANVRQAEGLDEGGAAMAVEPTQIOLSKRWE
TESKPAQRRNLFAQIHTVREEMVMTDMPLEI SVVPSTYLTETIHFVSOALLEVBO
LINALDCAKPEOLFKOESLKNIKOSLOOSSGRI DI IHSKTAALQOSATPVERVKI
QEARLSQDFQKYNKMYKDRQGRDRSVEKMRPFYDILKFNQMLTEAFQFLKTLQI
PENMHAKYKWLKQIGICQRTVVTNLNATGEEIIQOSSKIDASILQEKLSLNL
RQVEVCKLSRDKRLEKQNLSEFQORDNEFVLWEADNTASIPLEPKQOOLKE
KLEQVLEVEPLERQIGLKNETGFPVLVSAPISPEBQDKLENKILKQNLQWIKVS
RALPEKQIEIQAQILQOLSKLELDEPOLHILLMLSPRLNLEIYNQNOBGPFD
VKETEIAVOAKOPDVEELISKGQHLXKEKPAQVPRKLELDSLEWKAANRLIQEILRA
KOPDLAPGLTWIGASPTVTLVTPVVTKETI-SKLEMPSSLVLEPALADFNRAWT
ELTDHLSLLDQVLSQORVWNGLEDINEMIKOKATWQDLEQRPOLPEELITAAONLK
NKTISNOEARTITIDRIKIQONDEVEQHLQNRQKNQNLKMDSTQWLEAKEABQVL
GOARAKLESKEGPPYVDIAIKKJITETKQALDRQWQVNDVANDLALKILLRYSAD
DTRKYHMITENINASWRSIHKVSEREAALEETHLLOQFPDLDEKFLAMLTEATTA
NVLODATERKILLDSKVKELMKQMDLOQIEI EANTDVYHNLDSOKILRSLEGS
DAVLLQRLDNNFWKSLRKLSNIRSHLEASSDQWKRHLHLSQELLVWLQKDDSL
SROAPIGDGFPAQONDVHRAFKELKTEKPEVIMSTLETVRI FLTPOLLEGLEKLYO
EPRELPRERAQNTRLRKQAEVNTWPKLNLHSDMORKIDETLERLOELQLEAD
ELDLKLRQAEVTKGWSQVQVGLLIDSLDQHLKELKVALKEGEIAPLKENVSHVNDLQRL
TTGLQSLFYNLSTLELNTWKLLQVAVEDRVQLHEHRDPFASQHFLLSTS VQGP
WERAI SPNKVPYIINETHQTTCDWHPKMTLEYQLADLNNVFSAFTAMKRLRQKA
LCLDLSLSAADDALDNLKQNDQMDIQLINCLTITYDLRQEHNNLVNPLCYVD
MCLNWLNAVYDTRGRVRLSFKTGIISLCKAHLDEKRYLFRQVASSGFCQDRLR
GILLHDSIQI PROLEGEVASFGSNI EPSVRSFOPANNKPEI EALFLDMMWRLEPQSM
VWLPVLRVAAETAKHQAQKNI CKECPIIGFRYSLSKHFNHYDIQCSQFPSPVRAVKH
KWHYPMVEYCTPTTSGEDVRDFAKVLKKNFKTKYFPAKPRMGVLPVQTVLEGNMET
PVTLLNFWPVSAPASSPOLSHDDTHRIEYSLRLAEMNSNGSYLNDSPNESID
DEHLIIQYCSQNSQDSLOSPSPAQILISLESEERGELEIRIILADILEENRNQOAY
DLKQOHEHKGISLPSPPEWMTSPQPRDAELIAZAKILRQHKHLEARMQOILEDH
NKQLESOLHRLQOLLEQPOAKVNGTIVVSPSTSLORSOSSQPMLLRVVGSQTSDEM
GREDILSPQDPTSTGLEWMEQLNNSFPSSRGRTNTPGKPMREDTM"

BASE COUNT 4135 a 2524 c 2876 g 2911 t

Query Match 100.0%; Score 1301; DB 9; Length 12446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGACTTTCCAGCAGTTTCAAGACGAGACGATGATACATAGGGCCTTCAAGAGGGAATTGAA	60
Db	8606	CGACTTTCCAGCAGTTTCAAGACGAGACGATGATACATAGGGCCTTCAAGAGGGAATTGAA	8665
QY	61	AACCTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTTCTGACAGACCA	120
Db	8666	AACCTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTTCTGACAGACCA	8725
QY	121	GCCTTTGGAGAGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCCTCTCGAGAGAG	180
Db	8726	GCCTTTGGAGAGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCCTCTCGAGAGAG	8785
QY	181	AGCCACAGATGTCTACCTCGCTTCTACGAAGCAGAGCTGAGGAGGTCAATATCTGAGTGGGA	240
Db	8786	AGCCACAGATGTCTACCTCGCTTCTACGAAGCAGAGCTGAGGAGGTCAATATCTGAGTGGGA	8845
QY	241	AAAATTGAACCTGCATCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAGACT	300
Db	8846	AAAATTGAACCTGCATCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAGACT	8905
QY	301	CCAGGAACTTCAAGAGCCACGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT	360
Db	8906	CCAGGAACTTCAAGAGCCACGATGAGCTGGACCTCAAGCTGCGCCCAAGCTGAGGTGAT	8965
QY	361	CAAGGAGTCTCGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCCAAGATCACTGCA	420
Db	8966	CAAGGAGTCTCGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCCAAGATCACTGCA	9025
QY	421	GAAAGTCAAGGCACTTCGAGGAGAAATTTGCGCTCTTGAAAGAGAACTGAGCCACGTCAA	480

Db	9026	GAAAGTCAAGGCACTTCGAGGAGAAATTTGCGCTCTTGAAAGAGAACTGAGCCACGTCAA	9085
QY	481	TGACCTTGTCTCCAGCTTACACATTTTGGGCAATTCAGCTCTCACCGTATATACTCAGCAC	540
Db	9086	TGACCTTGTCTCCAGCTTACACATTTTGGGCAATTCAGCTCTCACCGTATATACTCAGCAC	9145
QY	541	TCTGGAAGACCTTGAACACACAGATGGAAGCTTTTGCAGTGGCCGCTGCGAGGACCGAGTCAG	600
Db	9146	TCTGGAAGACCTTGAACACACAGATGGAAGCTTTTGCAGTGGCCGCTGCGAGGACCGAGTCAG	9205
QY	601	GCAGCTGCATGATGAGCCACAGAGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTC	660
Db	9206	GCAGCTGCATGATGAGCCACAGAGGACTTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTC	9265
QY	661	TGTCCAGGCTCCCTGGGAGAGACCATCTCTCGCAAAACAAAGTGCCTACTATATCAACCA	720
Db	9266	TGTCCAGGCTCCCTGGGAGAGACCATCTCTCGCAAAACAAAGTGCCTACTATATCAACCA	9325
QY	721	CGAGCTCAAAACAACTTGTCTGGGACATCCCAAAATGACAGAGCTCTACCACTCTTTAGC	780
Db	9326	CGAGCTCAAAACAACTTGTCTGGGACATCCCAAAATGACAGAGCTCTACCACTCTTTAGC	9385
QY	781	TGACCTGAATAATGTTCAGATTTCTCAGCTTATAGGACTGCATGAAACTCCGAAAGACTGCA	840
Db	9386	TGACCTGAATAATGTTCAGATTTCTCAGCTTATAGGACTGCATGAAACTCCGAAAGACTGCA	9445
QY	841	GAAAGCCCTTGTCTGGATCTCTTGAGCCCTGTCAGTGCATGTGATGCTTGGACCGA	900
Db	9446	GAAAGCCCTTGTCTGGATCTCTTGAGCCCTGTCAGTGCATGTGATGCTTGGACCGA	9505
QY	901	CAACCTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTTATTAATTTTGGACCA	960
Db	9506	CAACCTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTTATTAATTTTGGACCA	9565
QY	961	TATTTATGACCCGCTGGAGCAAGAGACAAATTTTGGTCAACCTCCCTCTCTCGCTGGA	1020
Db	9566	TATTTATGACCCGCTGGAGCAAGAGACAAATTTTGGTCAACCTCCCTCTCTCGCTGGA	9625
QY	1021	TATGTCTCTGAACTGCTGCTGAAATGATATCTTATGATCGGAGCAACAGAGGAGTCCGCT	1080
Db	9626	TATGTCTCTGAACTGCTGCTGAAATGATATCTTATGATCGGAGCAACAGAGGAGTCCGCT	9685
QY	1081	CTGTCTTTTAAACCTGGCATCATTTTCCCTCTGTGTAAAGCACATTTTGGAAAGACAAAT	1140
Db	9686	CTGTCTTTTAAACCTGGCATCATTTTCCCTCTGTGTAAAGCACATTTTGGAAAGACAAAT	9745
QY	1141	ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCCAGCGAGGCTGGCCCT	1200
Db	9746	ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCCAGCGAGGCTGGCCCT	9805
QY	1201	CTTCTGCAATGATTTCTATCCAAATTTCCAAAGACAGTGTGGGTGAAGTTTGCATCTCTTGGGG	1260
Db	9806	CTTCTGCAATGATTTCTATCCAAATTTCCAAAGACAGTGTGGGTGAAGTTTGCATCTCTTGGGG	9865
QY	1261	CAGTAAATTTGAGCCAAAGTTCGAGAGTCTTCCAAATTTG 1301	
Db	9866	CAGTAAATTTGAGCCAAAGTTCGAGAGTCTTCCAAATTTG 9906	

RESULT 9

AX409637	Sequence 2284 from Patent WO0229103.	DNA	linear	PAT 14-JUN-2002
LOCUS	AX409637	13957 bp		
DEFINITION	AX409637			
ACCESSION	AX409637.1	GI:21442342		
VERSION	AX409637.1			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
AUTHORS	Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.			

TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2284 11-APR-2002;
GENE LOGIC INC (US)

FEATURES Location/Qualifiers
source 1..13957

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"

BASE COUNT 4602 a 2781 c 3122 g 3452 t

ORIGIN

Query Match 100.0%; Score 1301; DB 6; Length 13957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTCAGAGGAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
DB 8716 CGACTTTCCAGCAGTTCAGAGGAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 8775
QY 61 AACTAAAGAACCTGTATATCATGAGTACTCTTGAGACTGTACGAAATTTCTGACAGAGCA 120
DB 8776 AACTAAAGAACCTGTATATCATGAGTACTCTTGAGACTGTACGAAATTTCTGACAGAGCA 8835
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCGAGAGCTGCCTCTTGGAGAG 180
DB 8836 GCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCGAGAGCTGCCTCTTGGAGAG 8895
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
DB 8896 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 8955
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT 300
DB 8956 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT 9015
QY 301 CCAGGAATCTCAAGAGGCCAGGATGAGCTTCAAGCTTCAAGCTGCGCAAGCTGAGTGTAT 360
DB 9016 CCAGGAATCTCAAGAGGCCAGGATGAGCTTCAAGCTTCAAGCTGCGCAAGCTGAGTGTAT 9075
QY 361 CAAGGATCTTGGCAGCCCTGGGCGATTCCTCATGACTCTCTCAAGATCACTGCA 420
DB 9076 CAAGGATCTTGGCAGCCCTGGGCGATTCCTCATGACTCTCTCAAGATCACTGCA 9135
QY 421 GAAAGTCAAGCCTTCAGAGGAGAAATGGCGCTCTGAAAGAGAAAGTCAAGCAGCTCAA 480
DB 9136 GAAAGTCAAGCCTTCAGAGGAGAAATGGCGCTCTGAAAGAGAAAGTCAAGCAGCTCAA 9195
QY 481 TGACCTTGTCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
DB 9196 TGACCTTGTCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 9255
QY 541 TCTGGAGAGCTGNAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGAGCCGAGTCA 600
DB 9256 TCTGGAGAGCTGNAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGAGCCGAGTCA 9315
QY 601 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGACCTTCTTCCACGTC 660
DB 9316 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGACCTTCTTCCACGTC 9375
QY 661 TGTCAGGGTCCCTTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTCTATATCAACCA 720
DB 9376 TGTCAGGGTCCCTTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTCTATATCAACCA 9435
QY 721 CGAGACTCAAAACACTTGTCTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
DB 9436 CGAGACTCAAAACACTTGTCTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 9495
QY 781 TGACCTGAATATGTGAGATCTTCAAGCTTATAGAGCTGCCATGAAGTCCGAGAGCTGCA 840
DB 9496 TGACCTGAATATGTGAGATCTTCAAGCTTATAGAGCTGCCATGAAGTCCGAGAGCTGCA 9555
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGTGTGATGCTGATGAGTGGGA 900

DB 9556 GAAGGCCCTTTGCTTGGATCTCTTGGAGCTGTAGCTGCATGTGATGCTTGGACCA 9615
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATTAATTAATGTTGACCA 960
DB 9616 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATTAATTAATGTTGACCA 9675
QY 961 TATTTATGACCGCTTGGAGAGCAAGAGCAACAATTTGGTCAACGCTCTCTCGTGGGA 1020
DB 9676 TATTTATGACCGCTTGGAGAGCAAGAGCAACAATTTGGTCAACGCTCTCTCGTGGGA 9735
QY 1021 TATGTCCTGAACTGGCTGCTGAATGTTATGATACGGAGCAAGAGGATCCCTGT 1080
DB 9736 TATGTCCTGAACTGGCTGCTGAATGTTATGATACGGAGCAAGAGGATCCCTGT 9795
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTACAG 1140
DB 9796 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTACAG 9855
QY 1141 ATACTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACAGCGCAGCTGGGCT 1200
DB 9856 ATACTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACAGCGCAGCTGGGCT 9915
QY 1201 CCTTCTGATCATTTCTATCCAAATTTCCAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 1260
DB 9916 CCTTCTGATCATTTCTATCCAAATTTCCAGACAGTTGGGTGAAGTTGCATCCTTTGGGG 9975
QY 1261 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAATTTG 1301
DB 9976 CAGTAACATTGAGCCAAAGTGTCCGAGCTGCTTCCAATTTG 10016

RESULT 10
AX538581
LOCUS AX538581
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Chamberlain J.S. and Harper S.Q.
AUTHORS Mini-dysprophic nucleic acid and peptide sequences
TITLE Patent: WO 0229056-A 1 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES Location/Qualifiers
source 1..13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 4602 a 2781 c 3122 g 3452 t

ORIGIN

Query Match 100.0%; Score 1301; DB 6; Length 13957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTCAGAGGAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
DB 8716 CGACTTTCCAGCAGTTCAGAGGAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 8775
QY 61 AACTAAAGAACCTGTATATCATGAGTACTCTTGAGACTGTACGAAATTTCTGACAGAGCA 120
DB 8776 AACTAAAGAACCTGTATATCATGAGTACTCTTGAGACTGTACGAAATTTCTGACAGAGCA 8835
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCGAGAGCTGCCTCTTGGAGAG 180
DB 8836 GCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCGAGAGCTGCCTCTTGGAGAG 8895
QY 181 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
DB 8896 AGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 8955
QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT 300
DB 8956 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT 9015
QY 301 CCAGGAATCTCAAGAGGCCAGGATGAGCTTCAAGCTTCAAGCTGCGCAAGCTGAGTGTAT 360
DB 9016 CCAGGAATCTCAAGAGGCCAGGATGAGCTTCAAGCTTCAAGCTGCGCAAGCTGAGTGTAT 9075
QY 361 CAAGGATCTTGGCAGCCCTGGGCGATTCCTCATGACTCTCTCAAGATCACTGCA 420
DB 9076 CAAGGATCTTGGCAGCCCTGGGCGATTCCTCATGACTCTCTCAAGATCACTGCA 9135
QY 421 GAAAGTCAAGCCTTCAGAGGAGAAATGGCGCTCTGAAAGAGAAAGTCAAGCAGCTCAA 480
DB 9136 GAAAGTCAAGCCTTCAGAGGAGAAATGGCGCTCTGAAAGAGAAAGTCAAGCAGCTCAA 9195
QY 481 TGACCTTGTCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 540
DB 9196 TGACCTTGTCTGCGCAGCTTACCACCTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCAC 9255
QY 541 TCTGGAGAGCTGNAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGAGCCGAGTCA 600
DB 9256 TCTGGAGAGCTGNAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGAGCCGAGTCA 9315
QY 601 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGACCTTCTTCCACGTC 660
DB 9316 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGACCTTCTTCCACGTC 9375
QY 661 TGTCAGGGTCCCTTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTCTATATCAACCA 720
DB 9376 TGTCAGGGTCCCTTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTCTATATCAACCA 9435
QY 721 CGAGACTCAAAACACTTGTCTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
DB 9436 CGAGACTCAAAACACTTGTCTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 9495
QY 781 TGACCTGAATATGTGAGATCTTCAAGCTTATAGAGCTGCCATGAAGTCCGAGAGCTGCA 840
DB 9496 TGACCTGAATATGTGAGATCTTCAAGCTTATAGAGCTGCCATGAAGTCCGAGAGCTGCA 9555
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGTGTGATGCTGATGAGTGGGA 240

ETVTTVTRBQILVKEAQBELPPPPQKKRQITVDSEIRKRLDVIDITELHSMITRSEA
VLOSPEFAIRKRGNSFDLKKVNAJEREKAEKFRILQDASRAOALVEOMVNEGVNA
DSIKQASOLNSWISFECOLLSERLAWLEQNNIIFAYNLOOLOLEWTTAAENLWIKQ
PTTPEPTAKXSOLKICKBENWRLSGLQPIERLKIQSALKEKGGPFLDADVAF
TNHFKQVSDQAREKELQITFDLPPMRTOETMSAIRTWVQSESTKISIPOLSVTDY
EIMEORLELOALQESLQOQGLYLSYTSKEMSKAPSEISRKYQSEFIEBIRGWK
KLSOLVHEQCKLEBQNNKRLKQNIQTLEKWMFAEVDFLEKEWPAFGDSILKQL
QKRLVLDTQTTQPSLNSVNEGQKIKNEAPEFASRLTELKELNTOWDHWKQOYV
ARKEALGKFKTVLSKQDISEMHWNTQAEVEYLERDPYKTPDLOKAVEEMKRAK
BEAQKEKVKLLTESVNSVIAQAPVAQALKELETTITNYOMLCTRLNGKCKTLE
EYVACHWELLSYLEKANKWLENEFKLKTENIPGGAESISVLSLENLMSHSDNP
NQIRLAQITDQWDELINSELETFSNRRELHSEAVRQKLLEQSSQSQBQETPS
LHUIQSLTFIDQOLAYIADKYDAQMPQOAKIOSDITSHISEIUSEMKHNGKREA
AQRVLSIDVAQKLLQDSNKFPLQKPFANFELRQESKMLIDEVXOHLPALETKSVE
QEVVQSOLHCVNLSVSEKSEVEMVITGRQIVQKQTEPNKEDERVATLKLHY
NELGAKYTERKQOLEKCLSRKMRKEMNLTWMLAATDMELTKRGAECMPSNLSDE
VANGKATOKIEIKQVHLKSI TEVGEALKTVLGKETLVEDKLSLANSWIAVTSRAE
EWNLLLEYKHEWETPDQNDVHTIKLIOADTLLDESEKKPKQOQEDLRLKASLND
IRPKVDSRQANLWNRGDHCKLVEPOI SELNHRFAAISHRITKGRASIPKLELE
QFNSDIQKLEPLEAEIQOQVNLBEBDFKONNENEGTVKELLQEGDNLQRIIDER
KRBEIKIQOGLQTNALKDLRSORRKALETSHQWYQYKQADOLLKCLDIDIKKL
ASLPEPRDERKIKEIDRLOKKEELNAVERQAEGLSEDAAMAFTQIOLSKWRE
IESKPAFLRQINPAQILVREETMMVMTEDMPLEISYVSPSTYLTETTHVSQALLVEQ
LLEAPDLCAQDFDLQKESLAKNIKDSLOQSSGRIDIHSKKTALQASATPVERVK
QELASQLDPQWKVNMKQDRQDRPDRSEKWRPHYDIKIFNQMILTEAEQFLRKQI
FENWEHAKYKWLKLEQDQIGQOTVVRNLNATGEEIIQOOSKTIDASILQEKLSLNL
RWBEVCKQLSDRKRLEBQKILSEFORDLNEFVLEADNITASIPLEPGKEQOLKE
KLQVLLVEPLRQILQOLNETGQPVLSAPISPEQDKLENLQKTNLQWIKVXS
RALPEKQGETEAOIKDQGLEKLELLEBQNLHLLWLSPIRNLQYVNPQNBQFPD
VORTEIAVQAKQDPVEBLSKGLVKEKPAQVVKRKLLEDLSSEWKAVERNLLQBLA
KQPLAPLQPTIGASPTQVTLVTPQVTKETAISKLEMPSSILMLVPAALDNFAWT
ELDMLSLLOQVLSQWVMDLEDINEMIIKQKATWQDLEQRRPQLELITAAQNLK
NKYSQPARTIITDRIQONQWDEQHLQNRQONEMLKDSTOWLAKSEAEQVSL
QAPAKLESKEQPTVDAITOKKITETQLAKDLQWQNDVNDVANDLALKLRLDSYAL
DTRVHMRTENINASWRSIHKYSREAALEETHRLQFPPLDLEFLAWLTEAETTA
NVLDQATRKRLLEDSDGKVKLMKOWDQGEIEAHTDVTYVHNDENSQKILASLBSGD
DAVLQRLDNLMAKFWELSKLSNIRSHLEASSDQWKRHLHLSLOBLVWLOKDEL
SROAPIGGDPFVAKQNDVHEAKRELKQKPEVIMSTLTVRI FLEQPLEGLKLEY
EPRELPEBEAQNVTLLRQAEVENTWEKMLNHLADWQRKIDETLERLQBLQBATD
ELDLKQAEVITKSGVNDGLDILSDHLEKVKALRGELIAPKENVSHVNDLRAQOL
TTIGQISPVNLTDELNTLRKLLQVAVEDVRQLEHARDPQASQHLSTSVQGP
WEAISNKVPIYINETHOTTDHPKMTBELYQSLADANNVFP SAYTAMKLRLLQKA
LCIDLILSILSACADLQNDQNDQMDILQINCLTIVIVDRLEQBNHNLVNVPLQVD
MLNMLNVYDGTGRTRIVLSFKTIGIISLCKAHLQKVKYLPFKOVASSTGFCDDRL
GLLHDSIOIIPROLGVEVASPGNSIIPSVRSQFANKPEIEALFLDMWLEPQSM
VWLFVLRVAAETAKQAQKNCIKCEPILIGFRYLSLKHFNVDICQSCFFSRRVAKGH
KMHYPVVEYCTPTTSGEDVDRAKVLKKNKFRTRYIPAKPHRMGILPVQVLEGDMMET
PVTILNFWPVDSPASSPQSHDDTHSRIEHYASRLAEMENGSYLNDSISPNESID
DEHLIIHQYQCSLNQDSFQSPRSQAIIISLESBERGELERILADLEENLQABY
DLKQOHEKGLSPLPSPPEMPTSPQSPRDAELIAEAKILLRQHKLREARMQIILEDH
NKOLSELRLRLQLEPOAEAKVNGTIVTSRPSSTLSQRSDSSOPMLLRVVGSTSDSM
GEEDLLSPQDITSGLEEWEOQLNNSPPSRGRNTPGKPMREDTM"

BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 100.0%; Score 1301; DB 9; Length 13957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTTCCAGCAGTTGAGAGCAGACGATGATACATAGGCGCTTCAAGAGGAATTGAA 60
Db 8716 CGACTTCCAGCAGTTGAGAGCAGACGATGATACATAGGCGCTTCAAGAGGAATTGAA 8775
Qy 61 AACTAAAGACCTGTAATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
Db 8776 AACTAAAGACCTGTAATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 8835
Qy 121 GCCTTTGAGAGGATAGAGAAACTCTTACAGAGGCCAGAGAGCTGCTCTCTGAGAGAG 180
Db 8836 GCCTTTGAGAGGATAGAGAACTCTTACAGAGGCCAGAGAGCTGCTCTCTGAGAGAG 8895
Qy 181 AGCCAGAGATCTACTCGGCTTCTACGAAAGACAGGCTCAGAGGTCAATACTGAGTGGGA 240

Db 8896 AGCCAGAGATCTACTCGGCTTCTACGAAAGACAGGCTCAGAGGTCAATACTGAGTGGGA 8955
Qy 241 AAAATTGAACCTGCACCTCCGCTGACTGCGCAGAGAAAAATAGATCAGACCCCTTGAAAGACT 300
Db 8956 AAAATTGAACCTGCACCTCCGCTGACTGCGCAGAGAAAAATAGATCAGACCCCTTGAAAGACT 9015
Qy 301 CAGAACTTCAAGAGGCCACGAGATGAGCTGGACCTCAAGCTGCGCAGAACGCTGAGGTGAT 360
Db 9016 CAGAACTTCAAGAGGCCACGAGATGAGCTGGACCTCAAGCTGCGCAGAACGCTGAGGTGAT 9075
Qy 361 CAAGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCTCAAGATCACTCTGA 420
Db 9076 CAAGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCTCAAGATCACTCTGA 9135
Qy 421 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTCGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA 480
Db 9136 GAAAGTCAAGGCACCTTCGAGGAGAAAAATTCGCCCTCTGAAAGAGAACGCTGAGCCACGTCAA 9195
Qy 481 TGACCTTGCTCGCCAGCTTACACCTTTTGGGCAATTCAGCTCTCAGCGTATTAACCTCAGCAC 540
Db 9196 TGACCTTGCTCGCCAGCTTACACCTTTTGGGCAATTCAGCTCTCAGCGTATTAACCTCAGCAC 9255
Qy 541 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTCGAGGTGGCCGCTCGAGGACCGAGTCA 600
Db 9256 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTCGAGGTGGCCGCTCGAGGACCGAGTCA 9315
Qy 601 CGAGCTGATGAAGCCACACAGGACCTTTTGTTCAGCATCTCAGACACTTTCTTTCACAGTC 660
Db 9316 CGAGCTGATGAAGCCACACAGGACCTTTTGTTCAGCATCTCAGACACTTTCTTTCACAGTC 9375
Qy 661 TGTCCAGGGTCCCTGGGAGAGAGCCTCTGCCAAAACAAAGTGCCTACTATATCAACCA 720
Db 9376 TGTCCAGGGTCCCTGGGAGAGAGCCTCTGCCAAAACAAAGTGCCTACTATATCAACCA 9435
Qy 721 CGAGCTCAAAACAACTTGTCTGGGACCAATCCCAAATGACAGAGCTCTACAGTCTTTAGC 780
Db 9436 CGAGCTCAAAACAACTTGTCTGGGACCAATCCCAAATGACAGAGCTCTACAGTCTTTAGC 9495
Qy 781 TGACCTGATTAATGTCCAGATCTCAGCTTATAGGACTGCCATGAACTCCGAGACATGCA 840
Db 9496 TGACCTGATTAATGTCCAGATCTCAGCTTATAGGACTGCCATGAACTCCGAGACATGCA 9555
Qy 841 GAAGGCCCTTTGCTTGGATCTCTTGGCCTGTGAGCTGCAATGATGCTCTCTCGCTGGA 900
Db 9556 GAAGGCCCTTTGCTTGGATCTCTTGGCCTGTGAGCTGCAATGATGCTCTCTCGCTGGA 9615
Qy 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGCAGATTAATTAATTTGTTGACAC 960
Db 9616 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGCAGATTAATTAATTTGTTGACAC 9675
Qy 961 TATTTATCAGCGCTTGGAGCAAGACCAACAAATTTGTGTCAACGCTCCCTCTCTCGCTGGA 1020
Db 9676 TATTTATCAGCGCTTGGAGCAAGACCAACAAATTTGTGTCAACGCTCCCTCTCTCGCTGGA 9735
Qy 1021 TATGTGCTGAGCTGGCTGCTGAATGTTTATGATACGGGACGACAGGAGGATCCGCTG 1080
Db 9736 TATGTGCTGAGCTGGCTGCTGAATGTTTATGATACGGGACGACAGGAGGATCCGCTG 9795
Qy 1081 CCGTGTCTTTTAAACTGGCATCATTTCCCTCTGTGTAAGCACATTTTGNAGACAGTACAG 1140
Db 9796 CCGTGTCTTTTAAACTGGCATCATTTCCCTCTGTGTAAGCACATTTTGNAGACAGTACAG 9855
Qy 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGGCTGGGCT 1200
Db 9856 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGGCTGGGCT 9915
Qy 1201 CTTCTGATGATTTCTATCCAAATTCAGACAGTTGGGTGAGTTGATCTCTTTGGGG 1260
Db 9916 CTTCTGATGATTTCTATCCAAATTCAGACAGTTGGGTGAGTTGATCTCTTTGGGG 9975
Qy 1261 CAGTAACATTGAGCCAGGTGTCGAGAGCTGCTCCAAATTTG 1301
Db 9976 CAGTAACATTGAGCCAGGTGTCGAGAGCTGCTCCAAATTTG 10016


```

RESULT 12
AR220819      13977 bp      DNA      linear      PAT 26-SEP-2002
LOCUS
DEFINITION    Sequence 60 from patent US 6426186.
ACCESSION    AR220819
VERSION      AR220819.1 GI:23327696
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 13977)
AUTHORS      Jones, K.A., Volkmut, W. and Walker, M.G.
TITLE        Bone remodeling genes
JOURNAL      Patent: US 6426186-A 60 30-JUL-2002;
              Location/Qualifiers
FEATURES
source
1..13977
/organism="unknown"
BASE COUNT   4596 a 2765 c 3120 g 3453 t      43 others
ORIGIN
Query Match      99.2%; Score 1290; DB 6; Length 13977;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY      1 CGATTTCCAGGAGTTTCAGAGCAGAGACCATGTACATAGGGCTTCAGAGGGAATTCAA 60
DB      8716 CGACTTTCAGAGCAGTTTCAGAGCAGAGACCATGTACATAGGGCTTCAGAGGGAATTCAA 8775
QY      61 AACTAAGAACCTGTAATCATGAGTACTCTTGAGAGCTGTACGAATATTCTTGACAGAGCA 120
DB      8776 AACTAAGAACCTGTAATCATGAGTACTCTTGAGAGCTGTACGAATATTCTTGACAGAGCA 8835
QY      121 GCCTTTGGAAGAGCTAGAGAACTCTACAGAGGCCAGAGAGCTCCCTCGAGAGAG 180
DB      8836 GCCTTTGGAAGAGCTAGAGAACTCTACAGAGGCCAGAGAGCTCCCTCGAGAGAG 8895
QY      181 AGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATCTGAGTGGGA 240
DB      8896 AGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATCTGAGTGGGA 8955
QY      241 AAAATTGAACCTGACATCCCTGACCTGGCGAGAGAAAATAGATGAGACCCCTGAAAGACT 300
DB      8956 AAAATTGAACCTGACATCCCTGACCTGGCGAGAGAAAATAGATGAGACCCCTGAAAGACT 9015
QY      301 CCAGGAACTTCAGAGGCCAGGATGAGCTGAGCTCAAGCTCGCCCAAGCTGAGGTGAT 360
DB      9016 CCAGGAACTTCAGAGGCCAGGATGAGCTGAGCTCAAGCTCGCCCAAGCTGAGGTGAT 9075
QY      361 CAAGGGATCTGGCAGCCCTGGGCGATCTCTCATGACTCTCTCCAAGATCACTCGA 420
DB      9076 CAAGGGATCTGGCAGCCCTGGGCGATCTCTCATGACTCTCTCCAAGATCACTCGA 9135
QY      421 GAAGTCAAGGCACTTCAGAGGAAATTCGCTCTGAAAGAGAACTGAGCCACCTCAA 480
DB      9136 GAAAGTCAAGGCACTTCAGAGGAAATTCGCTCTGAAAGAGAACTGAGCCACCTCAA 9195
QY      481 TGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTTCAGCTCTCACCGTATAAACCCTCAGCAC 540
DB      9196 TGACCTTGTCTGCGCAGCTTACCACTTTGGGCATTTCAGCTCTCACCGTATAAACCCTCAGCAC 9255
QY      541 TCTGGAAGACTGGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGAGCCAGCTCAG 600
DB      9256 TCTGGAAGACTGGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGAGCCAGCTCAG 9315
QY      601 GCAGCTGATGAAGCCACAGGAGCTTGTGTCAGAGCTCTCAGCAGCTTCTTTTCCACGTC 660
DB      9316 GCAGCTGATGAAGCCACAGGAGCTTGTGTCAGAGCTCTCAGCAGCTTCTTTTCCACGTC 9375
QY      661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCCGCAAAACAAAGTGCCTTACTATATCAACCA 720
DB      9376 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCCGCAAAACAAAGTGCCTTACTATATCAACCA 9435

```

```

QY      721 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTTACCACTCTTACG 780
DB      9436 CGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTTACCACTCTTACG 9495
QY      781 TGACCTGAAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
DB      9496 TGACCTGAAATATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 9555
QY      841 GAAGGCCCTTTGCTGGATCTCTTGGAGCTGTGAGCTGTGATGATGATGATGATGATGATGATG 900
DB      9556 GAAGGCCCTTTGCTGGATCTCTTGGAGCTGTGAGCTGTGATGATGATGATGATGATGATGATG 9615
QY      901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCCTGACAGATTATTAATTTGATGACCA 960
DB      9616 CAACCTCAAGCAAAATGACAGCCCATGGATATCCCTGACAGATTATTAATTTGATGACCA 9675
QY      961 TATTTATGACCGGCTGGAGCAAGAGACACAACTTTGGTCAACGTCCTCTCTGCTGCTGGA 1020
DB      9676 TATTTATGACCGGCTGGAGCAAGAGACACAACTTTGGTCAACGTCCTCTCTGCTGCTGGA 9735
QY      1021 TATGTGTCTGAACCTGCTGCTGATCTTTATGATACGGAGCAAGAGGAGGATCCGCTGT 1080
DB      9736 TATGTGTCTGAACCTGCTGCTGATCTTTATGATACGGAGCAAGAGGAGGATCCGCTGT 9795
QY      1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAACTACAG 1140
DB      9796 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAACTACAG 9855
QY      1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAAGAGATTTTGTGACCAAGGAGCTCGGCC 1199
DB      9856 ATACCTTTTCAAGCAAGTGGCAAGTTTCAAGAGATTTTGTGACCAAGGAGCTCGGCC 9915
QY      1200 TCCTTTCGATGATTTCTATCCAAATTCAGAGCAGTTGGTGAAGTTGATCCTCTTTGGGG 1259
DB      9916 TCCTTTCGATGATTTCTATCCAAATTCAGAGCAGTTGGTGAAGTTGATCCTCTTTGGGG 9975
QY      1260 GCAGTAACANTGAGCCCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
DB      9976 GCAGTAACANTGAGCCCAAGTGTCCGAGCTGCTTCCAAATTTG 10017

```

```

RESULT 13
AF070485      13887 bp      mRNA      linear      MAM 09-DEC-1998
LOCUS
DEFINITION    Canis familiaris dystrophin mRNA, complete cds.
ACCESSION    AF070485
VERSION      AF070485.1 GI:3982750
KEYWORDS
SOURCE
ORGANISM
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 13887)
AUTHORS      Carville, K.S., Mann, C.J., Schatzberg, S.J. and Wilton, S.D.
TITLE        Direct Submission
JOURNAL      Submitted (04-JUN-1998) ANRI, Pathology, University of Western
              Australia, Verdun Street, Nedlands, WA 6018, Australia
FEATURES
source
1..13887
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="breed Golden retriever"
/db_xref="taxon:9615"
<1..289
290..11332
/codon_start=1
/product="dystrophin"
/protein_id="AAC83646.1"
/db_xref="GI:3982751"
/translation="MLWEEVEDCYBEEDYOKKFTKWNNAOFKSGKHIEHNFSL
QGRRLDLDLEGLTGQKPKPKGSTRHANNKALRVLQKNVDLVNIGSTDIVDG
NHKLTGLIWNIIHLHWQKVMKNIIMAGLQQTNSKILLSWVRQSTRNYFQVNVINPT

```

TSRSDGLNALNLIHSHRPOLFDWNSVVCQSATORLEHAFNIAKYQLGIEKLLDPEDY
ATTYPPKKSIMYITSLFOVLPOVSEIAIOEVMPLRPSOVRTREHFOIHHOMYSO
QITVSLAQGVRRPSPKPRKSYAYTOAYVTTSDPTSPPLPSHOLTEPDEKSGRS
LHTEANLQSVCTALBEVLSWLLSADALQAQGEISNDVEVEKEOPHHEGYNMGLTS
HQRGNVLQSGDLQITGKLSDEBTEVEQENLNSRWELRLRVASMEKQSNLKHVL
MDLQONKELNQLITKERTKMKPEKPLGPDIEDLKQVQOQHUKJEDLEQEQVRV
NSLTHMVVDESSEGHATIALEBQKLVQDRWANI CRWTEWRWLLDILLKWRFT
EEQLFSLASLEDAVNKIHTTFKQDSEVLNKLQALVKTDLEKKQOTWDKLCSL
NQDLSALNTVVAHMEAWLNDFAQRWNLVQKLEKSAQISOAVTTTPSITOTV
METVWTVTTHVHQAQBELPPPPKQKQIIVDSERKLDVDTIELHSHITRSE
AVLQSPFALVRKGNFSLCKVNAIEREKAKPKLODASRSQAALVEOMVNGVN
ADSIKQASQELNSRWIEFCOLLERLNLWLEYQNNLIITFNQLOQLBOMTTIABNLKT
QPTITSEPTAKGLKICDKEINRLSALQPIERLKIQSLAKEKGGPFLDADFA
FTNHQVFDVQVQKELQITSLDPSLPMRYOETWSTILTWTIQQSESTIPQVITE
YDIMEORGLQALQSLQOQNLNLSLTVKMSKAPLSDISRKYSOEFEEIEGR
WKLSSQLVHCQKLEQAKLKIQNH:KTLKWTIETDVFLEKEPALGDSLEIKR
QKQKLLVNDIOTIOPSINNVGAKQKMAEPFAPAGRLTELRELNTOWMYCRO
VYARKEALGGLDKTVSLQKDSMEHWMTQAEVEYLERFVKYPIDELQTAVERMKR
AKESAQKAKVLLVESVNSVLAQAPPAQKALKKELDTLTNTYNQWICRLNGKKT
LEEVACWHELLSHKSNKWLSEVEVKLTNTENISGABEEIAEVLDSLENLMQSHSD
NPQIRIATLTDGGVMDLEINBELTFNSRWELEHAEVRQKLEQSIQSAQIE
KSLHLIOESLSSDKQLAAYIADVDAAQMPQEAQIOSDLTSHESLEEMKHNQK
ETAQVLSQIDVACQKQVSMKFLRFQKPAFPEQRLQESKMLDLVKKMLPALETS
VEGEVQSOLNVCNLYKSLSEVKSEVEMVITGRQIVOKKOTENPKELDERVATLK
HYNELGAKTERKQLEKCLKLSKPMKEMNALTEWLAATDMELKXSASVEGMPNLD
SEVANGKATQKEIKQKVHLKSVTEVGEALKTVLGHKEMLVEDKLSLNSWIAVTSR
AEBWNLLEYQKHMETDQNDVYITNWIIOADLDESEKKPOQKEDILKRLKAEW
NDRPKVDSRQDAQANLNRDCHKRVVEPKISELHNFPAALSHIRKTKGASIPKE
LEQFNSDIQKLEPFAETQQGVNKEEDFNKMDSENEGTVKELLQRGNLQORITD
ERREBEIKIQOOLLQKHNALDRLSRRKKALETSHOWYQVKKQADDLKCLDSK
KLASLEPRDERIKSIEDLOKKEELNARVRAEGLSEDDGAAMAVEPTQIOLSKW
RETESKAPORRLNFAQIHTHVESVAVTEMDPELISVVPSTYLTELTHVSOALSEV
ZEILNAPLCAQDFEQLFKQESLKNIKDSLOQISORIDIENKTAALHSAITPRA
KLQELRLDFQERNVNNMTQGRFDSVKRRFHYDKMLQWLEAEQFLLKT
QIPENWEHAKYKRLKELQDQIGQORSVVRNATGEEIIQOSSKTASILQKLGSL
NLQWVECKQALRKEKQNLILSEFQDVNEFVLEEDANDVANIPIEPENQOOL
KEKLEOVKLAEELPLRQGLIKQNETGTVLVSAPLSEPODKLENKLOTNLOWIK
VSNRLPKQKEIEBAHVQHLQLEEQNLHLLMLSPIRNOLELYNOPNOTGPPDIKEIE
VAQAKQPDVEGILSKQHLXKPKATQPAKSKLEDLSMDKVVTLQOELRAKQPG
APGLTVTRAPQSVTLVTPQATKATSKJEMPSSLLEVPALADFNRAWTELIDW
LSLLDRVKSQRVWVGDLEINEMILOKATQDLEQRRPQLEELITAAQNJKNTSN
QEARTITDRIERIQOWBEVQHLQNRLOLTEMKLDSTOWLEAKEABOVLGQARA
KLESWEAPYVDAIQKITETQKALQROQINVDVANDLALKLRLDYSADDTKV
HMTENINASWAGIHKRLSEREALETHERLLQPPDLDEKFLAWTEATETANYQAL
ATHERLLEDSKVBELMKQNDQLOEIEBAHTDIDVNLNDENGOKVLSLESGDDAALL
ORLDDNMFKWBLKSLNIRSHLEASDDOKKHLHLSQELLVWLQIKDDELSQAP
IGDFFAVQKQNDVHRAFKRELKPKSPVIMSTLTVRIIFTEQPLEGLEKIQEPRKL
PPBERAQNVTRLRQKQAEVNTQWELNVSADWQKIDEALERLQELQEADELDDLK
LRQAEVKGWQVGLLIDSLQDHLKVKALRGEITPLKENVYVNDLARQLTTIGI
QLSPYNLTLDELTNRKLLQVAIEDIRIQLEAHRDFGAPSHFLSTVQGFWEARI
SPNKVPYINHEQTTCWDPKMTLEYQSLADLNVRFSAYRTAMKLRRLQKALCLDL
LSLSAACDALDQNLKQNDQPMQDLOVINCLTIIYDRLQEHNNLVNVPCLVDMCLNW
LNNVYDTRGRTRVLSPTGILISLCKAHLEDKYXLFKQVASSSTGFCQDORLGLLH
DSIQIPROGEVASFGSGNSIPRSCFPANNKPEIEAALFDWNRLEPQKSNWMLP
LHRVAAAEAKHQAKNICEKPIIGFRVSLKFNENYDICQSFPSGRVAKGHMHP
MVEYCTPTTSGEDVRDFAKVKNKTRKRYFAKPRMGYLPVQTVLEGNMETPTVLI
NFWPVSAPASPOLSHDDTHSRIEHVASRLKMNENGSYLNDSISPNESIDDSHLL
IOHWVDSLSPSPSPRPAOILISLESEERELERILADLEGRNRLQAEYDLKQ
OHEHKGSLSP
SQRHRLQLEQQAQKANGVITVSPFSLQSRSSSQMMLLRVVSQTSSEMGEDLI
LSPQDSTGLEVEVQNLNHSFPSSRGRNTPGKPMREDTM"
13333..513887
11789..11828
/rpt_type=tandem
/rpt_unit=ca
3'UTR repeat_region 4455 a 2906 c 3362 t
BASE COUNT 4455 a 2906 c 3362 t
ORIGIN

Query Match 88.7%; Score 1154.4; DB 4; Length 13887;

Best Local Similarity 93.0%; Pred. No. 0;

Matches 1209; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 2 GACTTTCCAGCAGTTCAGAGACAGACGATGTACATAGGCGCTTCAAGAGGAATTGAAA 61

8783 Db GACTTTCCAGCGTGCAGAGACAGAAATGATGTACACAGGCGCTTCAAGAGGGAATTGAAA 8842
62 QY ACTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTCTACGAAATATTTCTGACAGAGCAG 121
8843 Db ACGAAAGAACCTGTAAATCATGAGTACTCTTGAGACTCTACGAAATATTTCTGACAGAGCAG 8902
122 QY CTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAGA 181
8903 Db CTTTGAAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAGA 8962
182 QY GCCAGATGTCACTCGGCTTCTACGAAACAGAGCTGAGAGGTCATATCTGAGTGGGAA 241
8963 Db GCCAGATGTCACTCGGCTTCTACGAAACAGAGCTGAGAGGTCATATCTGAGTGGGAA 9022
242 QY AAATGAACCTGCACTCGGCTGAGTGGCAGAGAGAAATAGATGAGACCTTGAAGAGCTC 301
9023 Db AAATGAACCTGCACTCTGAGAGTGGCAGAGAGAAATAGAGAGGCGCTTGAAGAGCTC 9082
302 QY CAGGAACTTCAAGAGGCGCAGGATGAGCTGGAACCTCAAGCTGCCCAAGCTGAGGTGATC 361
9083 Db CAGGAGCTTCAAGAGGCGCAGGATGAGCTGGAACCTCAAGCTGCCCAAGCTGAGGTGATC 9142
362 QY AAGGATCTCTGGCAGCGCTGGGCGATCTCCTCATTGACTCTCTCCTCAAGATCACTCGAG 421
9143 Db AAGGATCTCTGGCAGCGCTGGGCGATCTCCTCATTGACTCTCTCCTCAAGATCACTCGAG 9202
422 QY AAAGTCAAGGCACTTTCAGAGAGAAATTTGGCGCTCTGAAAGAGAACTGAGGCGACGTCAT 481
9203 Db AAAGTCAAGGCGCTTTCAGAGAGAAATTTACACCTCTGAAAGAGAACTGAGGCGACGTCAT 9262
482 QY GACCTTCTGCCAGCTTACCACCTTGGCGATTCAGCTCTCACCCTGATATATCAACCTCAGCACT 541
9263 Db GACCTTCTGCCAGCTTACCACCTTGGCGATTCAGCTCTCACCCTGATATATCAACCTCAGCACT 9322
542 QY CTGGAAGACCTTGAAACACCACTGAGAGCTTCTGAGAGTGGCGCTGAGAGACCGAGTCAGG 601
9323 Db CTGGAAGACCTTGAAACACCACTGAGAGCTTCTGAGAGTGGCGCTGAGAGACCGAGTCAGG 9382
602 QY CAGCTGATGAGAGCGCAGAGGACTTTGGTCCAGCATCTCAGCACTTCTTCTTCCAGCTCT 661
9383 Db CAGCTGATGAGAGCGCAGAGGACTTTGGTCCAGCATCTCAGCACTTCTTCTTCCAGCTCT 9442
662 QY GTCCAGGCTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCGCTACTATATCAACAC 721
9443 Db GTCCAGGCTCCTCGGAGAGAGCCATCTCGCCAAACAAAGTGCGCTACTATATCAACAC 9502
722 QY GAGCTCAAACTTCTGCTGGGACCATCCCAAATGACAGAGCTTACAGTCTTTAGCT 781
9503 Db GAGCTCAAACTTCTGCTGGGACCATCCCAAATGACAGAGCTTACAGTCTTTAGCT 9562
782 QY GACCTGAATATGTACAGTCTTACAGCTTATAGGACTGCCCATGAACTCCGAGAGCTGCAG 841
9563 Db GACCTGAATATGTACAGTCTTACAGCTTATAGGACTGCCCATGAACTCCGAGAGCTGCAG 9622
842 QY AAGCCCTTTCTTGGATCTCTTGAGCTGTCAGTCTGATGCTGATGCTTGGACAGCAG 901
9623 Db AAGCCCTTTCTTGGATCTCTTGAGCTGTCAGTCTGATGCTGATGCTTGGACAGCAG 9682
902 QY AACTCAAGCAAAATGACAGCCCATGATATCTGAGATTTAATTTGTTGACCACT 961
9683 Db AACTCAAGCAAAATGACAGCCCATGATATCTGAGATTTAATTTGTTGACCACT 9742
962 QY ATTTATGATGCTGAGAGAGAGCACAACAAATTTGCTCAACCTCTCTCGGTGGAT 1021
9743 Db ATTTATGATGCTGAGAGAGAGCACAACAAATTTGCTCAACCTCTCTCGGTGGAT 9802
1022 QY ATGTGCTGAACCTGCTGCTGAATTTTATGACGAGAGAGAGAGAGGAGATCCCGTTC 1081
9803 Db ATGTGCTGAACCTGCTGCTGAATTTTATGACGAGAGAGAGAGGAGATCCCGGTC 9862
1082 QY CTGTCTTTAAACCTGCGCATCTTCCCTGCTGTAAGACATTTTGAAGACATTTGAAGACATGACAGA 1141

Db 9863 CTGCTCTTTAAACAGTCGCATCTTCTCTGTGTAAGGCCATTTTGAAGACCAAGTACAGA 9922
Qy 1142 TACCTTTTCAAGCAAGTGCAGATTTCACAGAGATTTCGTGACCCAGCGCAGCTGGGCTC 1201
Db 9923 TACCTCTTCAAGCAAGTGCAGATTTCGTGACCCAGCGCAGCTGGGCTC 9982
Qy 1202 CTTCTGCAATGATTCTATCCAAATTCACAGACAGTTTGGGTGAAGTTGCAATCTTTTGGGGGC 1261
Db 9983 CTTCTGCAATGATTCTATCCAGATCCCAAGACAGTTTGGGTGAAGTGCATCTGTCGGGGGC 10042
Qy 1262 AGTAACATTGAGCCAGAGTTCGGAGCTGCTTCCAAATTG 1301
Db 10043 AGTAACATTGAGCCAGAGTTCAGGAGCTGCTTCCAGTTTG 10082

RESULT 14
AX306153
LOCUS AX306153 13815 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 904 from Patent WO0188186.
ACCESSION AX306153
VERSION AX306153.1 GI:17645441
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188186-A 904 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
source
1. 13815
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 4577 a 2719 c 3025 g 3494 t
ORIGIN

Query Match 87.3%; Score 1135.2; DB 6; Length 13815;
Best local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 2 GACTTTCCAGAGTTCAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 61
Db 8708 GATTTCACAGAGTTCAGAGCAGACGATGTATACATAGGCGCTTCAAGAGGGAATTGAA 8767
Qy 62 ACTAAGAACTGTAATCATGAGTACTCTTGAAGACTGTAGCAATATTTCTGACAGAGCAG 121
Db 8768 ACTAAGAACTGTAATCATGAGTACTCTGAGAGCTGTGAGATATTTCTGACAGAGCAG 8827
Qy 122 CCTTTGAAGAGCTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTGAGGAGAGA 181
Db 8828 CCTTTGAAGAGCTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTGAGGAGAGA 8887
Qy 182 GCCCAGAACTGCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATCTAGTGGGAA 241
Db 8888 GCTCAGAACTGCACTCGGCTTCTACGAAAGCAGGCTGAAGAGGTCAACGCTGAATGGGAC 8947
Qy 242 AAATTGAACCTGCACTCGGCTGAGTGGCAGAGAAATAGATGAGACCTTGAAGACTC 301
Db 8948 AAATTGAACCTGCGCTCAGCTGATGGCAGAGAAATAGATGAGACTTGTGAAGACTC 9007
Qy 302 CAGGAACCTTCAAGAGGCGCAGGATGAGCTGAGACTCAAGCTGCGCCAGCAAGCTGAGTGATC 361
Db 9008 CAGGAACCTTCAAGAGGCTGCGGATGAATGCACTGCACTCAAGTTGGCCAAAGCTGAGGTGATC 9067
Qy 362 AAGGATCTCGGAGCCGCTGGGAGTCTCTCACTTGAATCTCTCAAGATCACTCTCGAG 421
Db 9068 AAGGATCTCGGAGCCAGTGGGGATCTCTCACTTGAATCTCTCAAGATCACTCTGAA 9127
Qy 422 AAAGTCAAGGCACTTCAGAGGAGAAATTGCGCTCTGAAAGAGAACTGAGCCACGTCAT 481

Db 9128 AAAGTCAAGGCACTTCGGGAGAAATTCACCTCTTAAAGAGATGTCAATCGTGCAAT 9187
Qy 482 GACCTTGTCTCGCGAGCTTACCACTTTTGGGCATTTCAAGCTCTCAACCGTATTAACCTCAGCACT 541
Db 9188 GACCTTGTCAATCAGCTGACACACTTGGGCATTTCAAGCTCTCAACCTTATAACCTCAGCACT 9247
Qy 542 CTGGAAGACTGTAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACGAGTCAAG 601
Db 9248 TTGGAAGACTGTAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGGAGCGTGTGAGA 9307
Qy 602 CAGCTGCATGAAGCCCAACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCT 661
Db 9308 CAGCTGCATGAAGCCCAACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCT 9367
Qy 662 GTCAGGCTCTCGGAGAGAGCCATCTCGCAACCAAAAGTGCCTTACTATATACACAC 721
Db 9368 GTTCAGGCTCTCGGAGAGAGCCATCTCACCACCAAAAGTGCCTTACTATATACACAC 9427
Qy 722 GAGACTCAAAACACTTGTCTGGGACCATCCAAATGACAGAGCTCTTACCAGTCTTTAGCT 781
Db 9428 GAGACCAAAACACTTGTCTGGGACCATCCCAAAAGTGCCTTACTATATACACAC 9487
Qy 782 GACCTGAATATCTCAGATTTCTCAGCTTATAGCACTGCGCATGAACTCCGAAGACTGCGAG 841
Db 9488 GACCTGAATATCTCAGATTTCTCGCGTATAGCACTGCCATGAAGCTCAGAAGGCTCCAG 9547
Qy 842 AAGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGCTGATGCTTGGACCAAGCACT 901
Db 9548 AAGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGCTGATGCTTGGACCAAGCACT 9607
Qy 902 AACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAAGATTAATTAATTTGTTGACCACT 961
Db 9608 AACCTCAAGCAAAATGACAGCCCATGGATATCTCTGCAAGATTAATTAATTTGTTGACCACT 9667
Qy 962 ATTATGACCGCTGGAGAGAGCAACAAATTTGCTCAAGCTCTCTCTCGCTGGAT 1021
Db 9668 ATTATGATCGTCTGGAGCAAGAGCAACAAATTTGCTCAAGCTCTCTCTCGCTGGAT 9727
Qy 1022 ATGTGCTGAACTGGCTGCTGAATGTTTATGATACGGAGCAACAGGAGGATCCGCTGTC 1081
Db 9728 ATGTGCTCAACTGGCTTCTCAATGTTTATGATACGGAGCAACAGGAGGATCCGCTGTC 9787
Qy 1082 CTGCTTTTAAACTGGCATCTTCCCTGTGTAAAGCACTTTGGAAGCAAGTACAGA 1141
Db 9788 CTGCTTTTAAACTGGCATCTTCTGTGTAAAGCACTTTGGAAGCAAGTACAGA 9847
Qy 1142 TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTC 1201
Db 9848 TACCTTTTCAAGCAAGTGGCAAGTTCAACGCTGCTTTGTGACAGCGTGGGCTCTT 9907
Qy 1202 CTTCTGATGATTTATCCAAATTCGAAGACAGTGGGTGAAGTGCATCTCTTTGGGGC 1261
Db 9908 CTTCTGATGATTTATCCAAATTCGAAGACAGTGGGTGAAGTGCATCTCTTTGGGGC 9967
Qy 1262 AGTAACATTGAGCCAGTGTCCGGAGCTGCTTCCAAATTG 1301
Db 9968 AGTAACATTGAGCCAGTGTCCGGAGCTGCTTCCAAATTG 10007

RESULT 15
AX338582
LOCUS AX338582 13815 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 2 from Patent WO0229056.
ACCESSION AX338582
VERSION AX338582.1 GI:25271088
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences

JOURNAL Patent: WO 0229056-A 2 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES
source 1. 13815
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 4577 a 2719 c 3025 g 3494 t
ORIGIN

Query Match 87.3%; Score 1135.2; DB 6; Length 13815;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 2 GACTTTCCAGAGCTTCAGAGCAAGATGATACATAGGCGCTTCAAGAGGGAATTGAAA 61
DB 8708 GATTTCCAGCAGTTCAGAGCAAGATGATACATAGGCGCTTCAAGAGGGAATTGAAA 8767
QY 62 ACTAAGAGACCTGTAATCATGACTACTCTTGAGACTGTACGAATATTCTGACAGAGCAG 121
DB 8768 ACTAAGAGACCTGTAATCATGACTACTCTTGAGACTGTGAGAAATATTCTGACAGAGCAG 8827
QY 122 CTTTGGAGAGCTAGAGAACTCTACAGAGCGCCAGAGAGCTGCTCCTGAGGAGAGA 181
DB 8828 CTTTGGAGAGCTAGAGAACTCTACAGAGCGCCAGAGAGCTGCTCCTGAGGAGAGA 8887
QY 182 GCCAGAAATGTCATCGGCTTCTACGAAAGCAGCTGAGAGGTCATATCTGAGTGGGAA 241
DB 8888 GCTCAGAAATGTCATCGGCTTCTACGAAAGCAGCTGAGAGGTCATATCTGAGTGGGAA 8947
QY 242 AAATTGAACCTGCACTCGCTGACTGCGCAGAGAAATAGATGAGACCTTTGAAAGACTC 301
DB 8948 AAATTGAACCTGCGCTGACTGCTGCGCAGAGAAATAGATGAGACCTTTGAAAGACTC 9007
QY 302 CAGGAACCTTCAAGAGGCGCAAGGATGAGCTGCAAGCTCAAGCTGCGCCCAAGCTGAGGTGATC 361
DB 9008 CAGGAACCTTCAAGAGGCGCAAGGATGAGCTGCAAGCTCAAGCTGCGCCCAAGCTGAGGTGATC 9067
QY 362 AAGGATCCTGGCAGCGCTGGCGGATCTCTCTATGCTCTCTCCAGATCACCCTGAG 421
DB 9068 AAGGATCCTGGCAGCGCTGGCGGATCTCTCTATGCTCTCTCCAGATCACCCTGAG 9127
QY 422 AAAGTCAAGGCACTTCGAGAGAGAAATGCGCTCTGAAAGAGAACGTGAGCCACGTCAT 481
DB 9128 AAAGTCAAGGCACTTCGAGAGAGAAATGCACTCTTAAAGAGATGTCATCGTGTCAT 9187
QY 482 GACCTTGTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACT 541
DB 9188 GACCTTGTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACT 9247
QY 542 CTGGAAGACCTGAAACACAGATGGAAGCTTCTGCAAGTGGCGCTGAGGACCGAGTCAGG 601
DB 9248 TTGGAGATCTGAAATACAGATGAGAGCTTCTACAGTGGCTGTGGAGGACCGGTGTCAGA 9307
QY 602 CAGCTGCAATGAAGCCCAACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCT 661
DB 9308 CAGCTGCAATGAAGCCCAACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCT 9367
QY 662 GTCCAGGCTCCCTGGGAGAGGCACTCTGCGCAACCAAGTGGCCCTACTATATCAACCC 721
DB 9368 GTTCAGGCTCCCTGGGAGAGGCACTCTCAGCAACCAAGTGGCCCTACTATATCAACCC 9427
QY 722 GAGACTCAAAACAACCTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCT 781
DB 9428 GAGACCCAAACCACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCT 9487
QY 782 GACCTGAATATGTCAGATTTCTAGCTTTATAGGACTGCCATGAATCTCCGAGAGCTGAG 841
DB 9488 GACCTGAATATGTCAGATTTCTAGCTTTATAGGACTGCCATGAATCTCCGAGAGCTGAG 9547
QY 842 AAGGCCCTTTGCTTGGATCTCTTGAGCTGTGAGCTGATGATGCTTTGGACGAGCAC 901
DB 9548 AAGGCCCTTTGCTTGGATCTCTTGAGCTGTGAGCTGATGATGCTTTGGACGAGCAC 9607

QY 902 AACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTAATTAATTTGTTGACCACT 961
DB 9608 AACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTAATTAATTTGTTGACTACA 9667
QY 962 ATTTATGACCGCTGGAGCAAGAGCAGCAACAATTTGTTCAAGTCCCTCTCTCGGTGGAT 1021
DB 9668 ATTTATGATCGTCTGGAGCAAGAGCAGCAACAATTTGTTCAATGTCCCTCTCTGTGTGGAT 9727
QY 1022 ATGTCTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTC 1081
DB 9728 ATGTCTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGTC 9787
QY 1082 CTGTCTTTTAAACTGGGCAATCATTTCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAGA 1141
DB 9788 CTGTCTTTTAAACTGGGCAATCATTTCTCTGTGTAAAGCACAATTTGGAAGACAAGTACAGA 9847
QY 1142 TACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACCGAGCGAGCTGGGCGTC 1201
DB 9848 TACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACCGAGCTGGGCGTC 9907
QY 1202 CTTTGTGATGATTTTATCCAAATTCAGAGCAGTTGGGTGAAGTTGATCCTTTGGGGGC 1261
DB 9908 CTTTGTGATGATTTTATCCAAATTCAGAGCAGTTGGGTGAAGTTGATCCTTTGGGGGC 9967
QY 1262 AGTAACATTTGAGCCAGTGTCCGGAGCTGCTTCCAATTTG 1301
DB 9968 AGTAACATTTGAGCCAGTGTCCGGAGCTGCTTCCAATTTG 10007

Search completed: February 1, 2004, 23:51:08
Job time : 4637.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 309.023 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100

Perfect score: 1301

Sequence: 1 cgactttccagcttcagc.....ccggagctgcttccaattg 1301

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_19Jun03.*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1301	100.0	1821	AAD37241	Human dystrophin r
2	1301	100.0	2169	AAD37232	Human dystrophin r
3	1301	100.0	3510	AAD37240	Human dystrophin m
4	1301	100.0	3531	AAD37238	Human dystrophin m
5	1301	100.0	3858	AAD37237	Human dystrophin m
6	1301	100.0	3999	AAD37234	Human dystrophin m
7	1301	100.0	4182	AAD37230	Human dystrophin m
8	1301	100.0	4476	AAD37259	Adeno-associated v

9	1301	100.0	4498	24	AAD37258	Adeno-associated v
10	1301	100.0	4825	24	AAD37257	Adeno-associated v
11	1301	100.0	4848	24	AAD37263	Adeno-associated v
12	1301	100.0	4966	24	AAD37256	Adeno-associated v
13	1301	100.0	4990	24	AAD37262	Adeno-associated v
14	1301	100.0	5060	24	AAD37264	Adeno-associated v
15	1301	100.0	5149	24	AAD37255	Adeno-associated v
16	1301	100.0	5339	24	ABX81998	DNA encoding mini-
17	1301	100.0	5462	24	ABX81999	DNA encoding mini-
18	1301	100.0	5952	22	AAD06794	Human dystrophin g
19	1301	100.0	8689	24	ABX82000	DNA encoding mini-
20	1301	100.0	11058	24	ABX37229	Human dystrophin p
21	1301	100.0	11241	24	ABX82005	CDNA encoding huma
22	1301	100.0	11443	24	ABX82002	DNA encoding mini-
23	1301	100.0	12923	10	AA90338	Sequence of human
24	1301	100.0	13957	24	ABT10904	Human breast cance
25	1301	100.0	13957	24	ABT10904	Human dystrophin g
26	1301	100.0	13957	24	ABX69900	Gene #2284 used to
27	1301	100.0	13957	24	ABX95786	CDNA encoding huma
28	1290	99.2	13977	24	ABX81959	Human bone remodel
29	1135.2	87.3	13815	24	ABX81960	CDNA encoding mous
30	1135.2	87.3	13815	24	ABX81960	Mouse ischaemic co
31	1135.2	87.3	19307	17	AAV18885	Shuttle vector pAd
32	1133.6	87.1	13815	19	AAV18885	Mus musculus dyst
33	1035.2	83.4	4402	21	AA248567	A rod shortened dy
34	1017.6	78.2	3446	24	AAD37242	Human dystrophin m
35	1017.6	78.2	4414	24	AAD37260	Adeno-associated v
36	1016.6	78.1	5417	24	ABX81997	DNA encoding mini-
37	1016	78.1	1434	24	AAD37243	Human dystrophin r
38	886	68.1	4402	21	AA248568	A rod shortened dy
39	750.4	57.7	4075	21	AA248569	A rod shortened dy
40	746.2	57.4	3747	21	AA248566	A rod shortened dy
41	692.6	53.2	3163	21	AA248571	A rod shortened dy
42	582.4	44.8	2005	25	ABT33376	NOVX DNA sequence
43	580.8	44.6	2463	25	ABT33375	NOVX DNA sequence
44	567.2	43.6	11096	24	ABX81962	CDNA encoding mous
45	553.8	42.6	6045	18	AAV74665	Utrophin minigene.

ALIGNMENTS

RESULT 1

AAD37241
ID AAD37241 standard; DNA; 1821 BP.

XX AC AAD37241;

XX DT 21-AUG-2002 (first entry)

XX DE Human dystrophin rod, hinge and CR domain regions encoding DNA #2.

XX KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.

OS Homo sapiens.

XX FN WO2000193695-A2.

XX PD 08-NOV-2001.

XX PF 27-APR-2001; 2001WO-USI3677.

XX PR 28-APR-2000; 2000US-200777P.

XX PA (XIAO/) XIAO X.

XX PI xiao X;

XX DR WPI; 2002-049342/06.

XX PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

Example 1: Page 52-53; 71pp; English.

XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin DNA fragment encoding rods R23 and R24, hinge H4 and CR
CC domain regions.

Sequence 1821 BP: 506 A; 451 C; 447 G; 417 T; 0 other; XX
SO

Query Match	100.0%	Score 1301;	DB 24;	Length 1821;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1301; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	CGAC	TTTCAG	CAGTTC	CGAG	CGAAG	CGATG	TACAT	AGG	CGCTT	CAAG	GAGGA	TTGAA	50
Db	103	CGAC	TTTCAG	CAGTTC	CGAG	CGAAG	CGATG	TACAT	AGG	CGCTT	CAAG	GAGGA	TTGAA	162
Qy	61	AACT	AAAGAA	CTGTAA	TCTCAT	GAGTACT	CTCTTG	AGACT	TGTAC	GAATAT	TTTCTG	CACAG	AGCA	120
Db	163	AACT	AAAGAA	CTGTAA	TCTCAT	GAGTACT	CTCTTG	AGACT	TGTAC	GAATAT	TTTCTG	CACAG	AGCA	222
Qy	121	GCCT	TTGGA	AGGACT	AGAGAA	AACTT	CTAC	CAGG	AGCC	CAGAG	AGCTGC	CTCTG	AGGAG	180
Db	223	GCCT	TTGGA	AGGACT	AGAGAA	AACTT	CTAC	CAGG	AGCC	CAGAG	AGCTGC	CTCTG	AGGAG	282
Qy	181	AGCC	CAGAAT	GTCACT	TCGCTT	CTAC	GAAAC	CGCGT	GTGAG	GGTCA	ATACT	GTAGT	TGGA	240
Db	283	AGCC	CAGAAT	GTCACT	TCGCTT	CTAC	GAAAC	CGCGT	GTGAG	GGTCA	ATACT	GTAGT	TGGA	342
Qy	241	AAAA	TTGAA	CCCTG	CACTCG	CTG	CGC	AGAG	AAAA	ATAG	TAG	TAG	AGAG	300
Db	343	AAAA	TTGAA	CCCTG	CACTCG	CTG	CGC	AGAG	AAAA	ATAG	TAG	TAG	AGAG	402
Qy	301	CCAG	GAAC	TTCA	AGAG	GCAC	GGATG	AGCT	TCG	ACCT	CAAG	CTGC	CAAGCT	360
Db	403	CCAG	GAAC	TTCA	AGAG	GCAC	GGATG	AGCT	TCG	ACCT	CAAG	CTGC	CAAGCT	462
Qy	361	CAAG	GGAT	CTCTG	CGAC	CCCGT	GGGCG	ATCT	CCCT	CATT	GACTCT	CTCCA	AGAT	420
Db	463	CAAG	GGAT	CTCTG	CGAC	CCCGT	GGGCG	ATCT	CCCT	CATT	GACTCT	CTCCA	AGAT	522
Qy	421	GAAG	CTCA	AGGC	ACCT	TCG	AGG	AGAA	TTTGG	CGCTCT	CGA	AGAG	AACTG	480
Db	523	GAAG	CTCA	AGGC	ACCT	TCG	AGG	AGAA	TTTGG	CGCTCT	CGA	AGAG	AACTG	582
Qy	481	TGAC	CTTGT	CTCG	CCAG	CTTAC	CACT	TTTGG	GCATTC	CAGCTCT	CAC	CGTAA	CCCT	540
Db	583	TGAC	CTTGT	CTCG	CCAG	CTTAC	CACT	TTTGG	GCATTC	CAGCTCT	CAC	CGTAA	CCCT	642
Qy	541	TTCT	GGA	AGAC	CTTGA	ACAC	AGAT	TGA	AGTCT	TG	CAG	GTG	CGT	600
Db	643	TTCT	GGA	AGAC	CTTGA	ACAC	AGAT	TGA	AGTCT	TG	CAG	GTG	CGT	702
Qy	601	GCAG	CTGC	ATGA	AGCC	CCAC	CAG	GGACT	TTTGGT	CTC	CAG	CATCT	CAG	660
Db	703	GCAG	CTGC	ATGA	AGCC	CCAC	CAG	GGACT	TTTGGT	CTC	CAG	CATCT	CAG	762
Qy	661	TGTC	CCAG	GGTCT	CTGG	AGAG	AGCC	ATCT	CGC	CA	AA	CAAG	TG	720
Db	763	TGTC	CCAG	GGTCT	CTGG	AGAG	AGCC	ATCT	CGC	CA	AA	CAAG	TG	822

rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -

Example 1; Page 45-46; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin DNA fragment encoding rods R22, R23 and R24, hinge H4 and CR domain regions.

Sequence 2169 BP; 623 A; 529 C; 524 G; 493 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 2169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTCCAGCAGTTCAGAGCAGACGATGACATAGGCGCTTCAGAGGAAATTGAA 60
451 CGACTTCCAGCAGTTCAGAGCAGACGATGACATAGGCGCTTCAGAGGAAATTGAA 510
61 AACTAAAGAACCTGTAAATCATAGTAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
511 AACTAAAGAACCTGTAAATCATAGTAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 570
121 GCCTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCCTGAGGAGAG 180
571 GCCTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCCTGAGGAGAG 630
181 AGCCCAAGATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGA 240
631 AGCCCAAGATGTCACCTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGA 690
241 AATATTGAACCTGACCTCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAGACT 300
691 AATATTGAACCTGACCTCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAGACT 750
301 CCAGGAACCTTCAAGAGCCACCGATGAGCTGGACCTCAAGCTGGCGCAAGCTGAGGTGAT 360
751 CCAGGAACCTTCAAGAGCCACCGATGAGCTGGACCTCAAGCTGGCGCAAGCTGAGGTGAT 810
361 CAAGGGATCTGGCAGCCCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 420
811 CAAGGGATCTGGCAGCCCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 870
421 GAAAGTCAGGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAGAGCTGAGCCACGTCGA 480
871 GAAAGTCAGGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAGAGCTGAGCCACGTCGA 930
481 TGACCTTGCTGCGCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGGTATAAAGCTCAGCAC 540
931 TGACCTTGCTGCGCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGGTATAAAGCTCAGCAC 990
541 TCTGGAAGACCTGAAACACAGATGGAAGCTTTCAGAGGTGGCGGTGAGGACCGAGTCTAG 600
991 TCTGGAAGACCTGAAACACAGATGGAAGCTTTCAGAGGTGGCGGTGAGGACCGAGTCTAG 1050
601 GCAGCTCATGAGCCACAGGAGCTTTGGTTCAGGATCTCAGCACTTTCTTTCCACGTC 660
1051 GCAGCTCATGAGCCACAGGAGCTTTGGTTCAGGATCTCAGCACTTTCTTTCCACGTC 1110
661 TGTCAGGGTCTCTGGGAGAGAGCACTCTCGCAAAACAAAGTGCCTACTATATCAACCA 720
1111 TGTCAGGGTCTCTGGGAGAGAGCACTCTCGCAAAACAAAGTGCCTACTATATCAACCA 1170
721 CGAGACTCAACAACTTGCTGGGAGACCATCCCAAAATGACAGAGCTCTTACCACTCTTTAGC 780

Db 1171 CGAGACTCAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTTACCACTCTTTAGC 1230
Qy 781 TGACCTGAATTAATGCTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAAGACTGCA 840
Db 1231 TGACCTGAATTAATGCTCAGATTCTCAGCTTATAGACTGCCATGAAACTCCGAAAGACTGCA 1290
Qy 841 GAAGGCCCTTTGCTTGGATCTCTTAGCCCTGTCAGCTGTCATGATGCCCTTGGACCAAGCA 900
Db 1291 GAAGGCCCTTTGCTTGGATCTCTTAGCCCTGTCAGCTGTCATGATGCCCTTGGACCAAGCA 1350
Qy 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCGAGATTATTAATTTGTGACCAAC 960
Db 1351 CAACCTCAAGCAAAATGACAGCCCATGGATATCCTGCGAGATTATTAATTTGTGACCAAC 1410
Qy 961 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGGA 1020
Db 1411 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGGA 1470
Qy 1021 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGGA 1080
Db 1471 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGGGA 1530
Qy 1081 CCTGCTTTTAAACTGGCATCATTTCCCTGTTGTAAGACACATTTGGAAGACAAGTACAG 1140
Db 1531 CCTGCTTTTAAACTGGCATCATTTCCCTGTTGTAAGACACATTTGGAAGACAAGTACAG 1590
Qy 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAAAGGATTTTGTGACAGCGAGGCTGGGCTCT 1200
Db 1591 ATACCTTTTCAAGCAAGTGGCAAGTTCAAAGGATTTTGTGACAGCGAGGCTGGGCTCT 1650
Qy 1201 CCTTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCACTCTTTGGGGG 1260
Db 1651 CCTTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGGTGAAGTTGCACTCTTTGGGGG 1710
Qy 1261 CAGTAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
Db 1711 CAGTAACATTGAGCAAGTGTCCGAGCTGCTTCCAAATTTG 1751

RESULT 3
AAD37240
ID AAD37240 standard; DNA; 3510 BP.
XX
AC AAD37240;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3510.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
WI 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -

XX Example 1; Page 51-52; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene is operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus, hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX SQ Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 3510;
 Best Local Similarity 100.0%; Pred No. 0;
 Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTCCAGCAGTTCAGAGCAGAACCGATGACATAGGGCTTCAAGAGGAATTGAA 60
 DB 1779 CGACTTCCAGCAGTTCAGAGCAGAACCGATGACATAGGGCTTCAAGAGGAATTGAA 1838

QY 61 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
 DB 1839 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 1898

QY 121 GCCTTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCTCGAGAGAG 180
 DB 1899 GCCTTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCTCGAGAGAG 1958

QY 181 AGCCAGAACTGCTACCTCGCTTCTAAGAGAGGCTGAGGAGTCAATCTGAGTGGGA 240
 DB 1959 AGCCAGAACTGCTACCTCGCTTCTAAGAGAGGCTGAGGAGTCAATCTGAGTGGGA 2018

QY 241 AAAATTGAACCTGCTACCTCGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 DB 2019 AAAATTGAACCTGCTACCTCGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2078

QY 301 CAGGAACTTCAAGAGCCACGGATGAGTGGAGCTCAAGAGCTGCGCAAGCTGAGGTGAT 360
 DB 2079 CAGGAACTTCAAGAGCCACGGATGAGTGGAGCTCAAGAGCTGCGCAAGCTGAGGTGAT 2138

QY 361 CAAGGGATCTGCGAGCCGCTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGA 420
 DB 2139 CAAGGGATCTGCGAGCCGCTGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGA 2198

QY 421 GAAAGTCAAGGCACTTCGAGAGAGAAATTTGGCTCTGAAAGAGAACTGAGCGCAAGTCAA 480
 DB 2199 GAAAGTCAAGGCACTTCGAGAGAGAAATTTGGCTCTGAAAGAGAACTGAGCGCAAGTCAA 2258

QY 481 TGACCTTCTGCGAGCTTACCACTTTGGCACTTTCAGCTCTCACCTATAAAGCTCAGCAC 540
 DB 2259 TGACCTTCTGCGAGCTTACCACTTTGGCACTTTCAGCTCTCACCTATAAAGCTCAGCAC 2318

QY 541 TCTGGAAGAGCTGAACACCAAGATGGAAGCTTCTGCGAGGTGGCGCTCGAGGACCGAGTCA 600
 DB 2319 TCTGGAAGAGCTGAACACCAAGATGGAAGCTTCTGCGAGGTGGCGCTCGAGGACCGAGTCA 2378

QY 601 GCAGCTGATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGCACTTTCTTCCACGTC 660
 DB 2379 GCAGCTGATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGCACTTTCTTCCACGTC 2438

QY 661 TGTCCAGGCTCCCTGGGAGAGAGCACTCTCCCAAAACAAAGTGCCTACTATATCAACCA 720
 DB 2439 TGTCCAGGCTCCCTGGGAGAGAGCACTCTCCCAAAACAAAGTGCCTACTATATCAACCA 2498

QY 721 CGAGACTCAAAACAACTTGTGGGAGACATCCCAAAATGACAGAGCTCTACCACTCTTAGC 780

DB 2499 CGAGACTCAAAACAACTTGTGGGAGACATCCCAAAATGACAGAGCTCTACCACTCTTAGC 2558

QY 781 TGACCTGAATAATGTCAGATCTCAGCTTATAGGACTGCATGAAACTCGAGAGCTGCA 840

DB 2559 TGACCTGAATAATGTCAGATCTCAGCTTATAGGACTGCATGAAACTCGAGAGCTGCA 2618

QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGTCATGTGATGTCCTTGGACAGCA 900

DB 2619 GAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAGCTGTCATGTGATGTCCTTGGACAGCA 2678

QY 901 CAACCTCAAGCAAAATGACAGCCCATGGAATCTCTGAGATTAATTAATTTTGACCA 960

DB 2679 CAACCTCAAGCAAAATGACAGCCCATGGAATCTCTGAGATTAATTAATTTTGACCA 2738

QY 961 TATTATGACCGCTGAGCAGAGCAGCAACAATTTGGTCAACCTGCTCTCTGCGTGA 1020

DB 2739 TATTATGACCGCTGAGCAGAGCAGCAACAATTTGGTCAACCTGCTCTCTGCGTGA 2798

QY 1021 TATGTGCTCAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080

DB 2799 TATGTGCTCAACTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 2858

QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGTAAGACAACTGTAAG 1140

DB 2859 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGTAAGACAACTGTAAG 2918

QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGAGATTTGTACCCAGCGCAGGCTGGGCT 1200

DB 2919 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGAGATTTGTACCCAGCGCAGGCTGGGCT 2978

QY 1201 CCTTCTGCATGATCTATCCAAATTCAGACAGATTTGGGTGAAGTTGCATCTTTCGGGG 1260

DB 2979 CCTTCTGCATGATCTATCCAAATTCAGACAGATTTGGGTGAAGTTGCATCTTTCGGGG 3038

QY 1261 CAGTAACATTGAGCGCAAGTGTCCGAGCTGCTTCCCAATTG 1301

DB 3039 CAGTAACATTGAGCGCAAGTGTCCGAGCTGCTTCCCAATTG 3079

RESULT 4
 AAD37238
 ID AAD37238 standard; DNA; 3531 BP.
 XX
 AC AAD37238;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE Human dystrophin minigene delta3531.
 XX
 KW Human; dyetrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200183695-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US13677.
 XX
 PR 28-APR-2000; 2000US-200777P.
 XX
 PA (XIAO/) XIAO X.
 XX
 PI Xiao X;
 XX
 DR WPI; 2002-049342/06.
 XX
 PT New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -

XX PS Example 1; Page 50-51; 71pp; English.

XX CC The present invention relates to an isolated nucleotide sequence encoding
XX CC a dystrophin minigene. The minigene comprises N-terminal or modified
XX CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX CC invention also relates to a recombinant adeno-associated virus (AAV)
XX CC comprising dystrophin minigene operably linked to an expression control
XX CC element. The dystrophin minigene in operable linkage with an expression
XX CC control element, in a recombinant adeno-associated virus or retrovirus is
XX CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX CC dystrophy (BMD) in a mammalian subject. The present sequence is human
XX CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
XX CC hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and
XX CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX SQ Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 3531;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60

DB 1800 CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 1859

QY 61 AACTAAAGAACCTGTATATCATGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120

DB 1860 AACTAAAGAACCTGTATATCATGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 1919

QY 121 GCCTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCTCCTGAGGAGAG 180

DB 1920 GCCTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCTCCTGAGGAGAG 1979

QY 181 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTCAGTGGGA 240

DB 1980 AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATACTCAGTGGGA 2039

QY 241 AAAATTGAACCTGCACTCCGCTGATCGGCGATTCCTCAATGACTCTCTCCAAAGATCACCTCGA 420

DB 2040 AAAATTGAACCTGCACTCCGCTGATCGGCGATTCCTCAATGACTCTCTCCAAAGATCACCTCGA 2099

QY 301 CCAGGAATCTCAAGAGGCCAGGATGAGTGGACCTCAAGCTCGCGCAAGCTGAGGTGAT 360

DB 2100 CCAGGAATCTCAAGAGGCCAGGATGAGTGGACCTCAAGCTCGCGCAAGCTGAGGTGAT 2159

QY 361 CAAGGGATTCCTGGCAGCCCGTGGCGGATTCCTCAATGACTCTCTCCAAAGATCACCTCGA 420

DB 2160 CAAGGGATTCCTGGCAGCCCGTGGCGGATTCCTCAATGACTCTCTCCAAAGATCACCTCGA 2219

QY 421 GAAAGTCAAGGCATTTCCAGAGGAATTTGGCTCTGAAAGAGAAAGTGGAGCCACGTCAA 480

DB 2220 GAAAGTCAAGGCATTTCCAGAGGAATTTGGCTCTGAAAGAGAAAGTGGAGCCACGTCAA 2279

QY 481 TGACTTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAATCTCAGCAC 540

DB 2280 TGACTTTGCTCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAATCTCAGCAC 2339

QY 541 TCTGGAAGACCTGAAACACAGATGGAAGCTTTGCGAGGTGGCGGTGAGGACCGAGTCAAG 600

DB 2340 TCTGGAAGACCTGAAACACAGATGGAAGCTTTGCGAGGTGGCGGTGAGGACCGAGTCAAG 2399

QY 601 GCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCATCTTCTTTTCCACGTC 660

DB 2400 GCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCATCTTCTTTTCCACGTC 2459

QY 661 TGTCAGGGTCCCTGGAGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCA 720

DB 2460 TGTCAGGGTCCCTGGAGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCA 2519

QY 721 CGAGACTCAACAACACTTGCTGGGAGCCATCCCAAAATGACAGAGCTCTACCACTTCCTTAGC 780

DB 2520 CGAGACTCAACAACACTTGCTGGGAGCCATCCCAAAATGACAGAGCTCTACCACTCTTAGC 2579

QY 781 TGACCTGAAATATCTCAGATCTCAGCTTATAGACTGCCATGAACTCCGAGAGCTGCA 840

DB 2580 TGACCTGAAATATCTCAGATCTCAGCTTATAGACTGCCATGAACTCCGAGAGCTGCA 2639

QY 841 GAAGGCCCTTTGGCTTGGATCTCTTGGAGCTCTCAGCTGCGATGTGATGCTTGGAGCCAGCA 900

DB 2640 GAAGGCCCTTTGGCTTGGATCTCTTGGAGCTCTCAGCTGCGATGTGATGCTTGGAGCCAGCA 2699

QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATTAATTTGTTGACCCAC 960

DB 2700 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATTAATTTGTTGACCCAC 2759

QY 961 TATTTATGACCGCTTGGAGCAAGAGCACAAATTTGGTCAACCTCTCCTCTCTCGGTGGA 1020

DB 2760 TATTTATGACCGCTTGGAGCAAGAGCACAAATTTGGTCAACCTCTCCTCTCTCGGTGGA 2819

QY 1021 TATGTGCTGAACTGGCTGCTGAAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 1080

DB 2820 TATGTGCTGAACTGGCTGCTGAAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGT 2879

QY 1081 CCTGTCTTTTAAACTGSCATCATTTCCCTGTGTAAGACACATTTGGAAGCAAGTACAG 1140

DB 2880 CCTGTCTTTTAAACTGSCATCATTTCCCTGTGTAAGACACATTTGGAAGCAAGTACAG 2939

QY 1141 ATACTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 1200

DB 2940 ATACTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 2999

QY 1201 CCTCTGCTGATCTTATCCAAATTCGAAGACAGTTCGGTGAAGTTGCATCCTTTGGGG 1260

DB 3000 CCTCTGCTGATCTTATCCAAATTCGAAGACAGTTCGGTGAAGTTGCATCCTTTGGGG 3059

QY 1261 CAGTAACTTCCAGCAAGTGTCCGGAGCTGCTTCCAAATTTG 1301

DB 3060 CAGTAACTTCCAGCAAGTGTCCGGAGCTGCTTCCAAATTTG 3100

RESULT 5

AAD37237

ID AAD37237 standard; DNA; 3858 BP.

XX AC AAD37237;

XX AC AAD37237;

DT 21-AUG-2002 (first entry)

XX Human dystrophin minigene delta3849.

DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX Homo sapiens.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-2007779.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

DQ

GGGAGTTCTTGTTCATCCCAACAACTCCTCGAAGGCCTTGCTCATCTTGA
780

FY

GAGACTCAAACAATACTTGTCTGGACCATTCCCAAAATGACAGAGCTCTACCA
790

PT New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT New dystrophin minigene for treating Duchenne or Becker
PT dystrophy comprises an N-terminal domain or modified N-
PT rod repeats, H1 and H4 domains and a cysteine rich domain

dystrophin gene -

Example 1; Page 46-47; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus, hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 3999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTTCCAGAGCTTCAGAGCAGACGATGTACATAGGCGCTTCAGAGGGAATTGAA 60
2268 CGACTTTCCAGAGCTTCAGAGCAGACGATGTACATAGGCGCTTCAGAGGGAATTGAA 2327

61 AACATAAGAACCTGTATCATGTAGTACTTGTAGACTGTACGAATATTTCTGACAGCA 120
2328 AACTAAGAACCTGTATCATGTAGTACTTGTAGACTGTACGAATATTTCTGACAGCA 2387

121 GCCTTTGGAAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTCCCTCTGAGGAGAG 180
2388 GCCTTTGGAAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTCCCTCTGAGGAGAG 2447

181 AGCCCAAGATGTCACTCGGCTTCTAGAAAGAGGCTGAGGAGTCAATACTGAGTGGCA 240
2448 AGCCCAAGATGTCACTCGGCTTCTAGAAAGAGGCTGAGGAGTCAATACTGAGTGGCA 2507

241 AAAATTGAACTGCACCTCCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAAGACT 300
2508 AAAATTGAACTGCACCTCCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAAGACT 2567

301 CCAGGAATCTCAGAGGCCAGGATGAGCTTGGACCTCAAGCTCGGCCCAAGCTGAGTGAT 360
2568 CCAGGAATCTCAGAGGCCAGGATGAGCTTGGACCTCAAGCTCGGCCCAAGCTGAGTGAT 2627

361 CAAGGGATCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAAGATCACCTCGA 420
2628 CAAGGGATCTGGCAGCCCGTGGCGGATCTCCTCATTTGACTCTCTCAAGATCACCTCGA 2687

421 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAACTGAGGCAACGTCGCA 480
2688 GAAAGTCAAGGCACTTCGAGGAGAAATTCGCTCTGAAAGAGAACTGAGGCAACGTCGCA 2747

481 TGACCTTGTCTGGCAGCTTACCACTTTGGGCACTTCAAGCTCTCACCGTATAAATCAGAC 540
2748 TGACCTTGTCTGGCAGCTTACCACTTTGGGCACTTCAAGCTCTCACCGTATAAATCAGAC 2807

541 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTCGAGGTGGCGCTCGAGGACCGAGTCAAG 600
2808 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTCGAGGTGGCGCTCGAGGACCGAGTCAAG 2867

601 GCAGCTGCATGAAGCCACAGGAGACTTTGCTCAGCATCTCAGCATCTTCTTTCACGTC 660
2868 GCAGCTGCATGAAGCCACAGGAGACTTTGCTCAGCATCTCAGCATCTTCTTTCACGTC 2927

661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCA 720
2928 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACCA 2987

721 CGAGACTCAAAACAACTTGTCTGGGAGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780

Db 2988 CGAGACTCAAAACAATCTGCTGGACCATCCCAAAATGACAGAGCTCTACCACTTTAGC 3047
QY 781 TGACCTGAATATGTGAGATTTCTCAGCTTATAGACTGCGCATGAAACTCCGAAAGCTGCA 840
Db 3048 TGACCTGAATATGTGAGATTTCTCAGCTTATAGACTGCGCATGAAACTCCGAAAGCTGCA 3107
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGCGATGTGATGCTTGGACCCAGCA 900
Db 3108 GAAGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGCGATGTGATGCTTGGACCCAGCA 3167
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCCAGATTTATTAATTTTGAACAC 960
Db 3168 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCCAGATTTATTAATTTTGAACAC 3227
QY 961 TATTATGACCGCTGCGAGCAAGAGCACAAATTTGGTCAACGCTCCCTCTCTCGGTGGA 1020
Db 3228 TATTATGACCGCTGCGAGCAAGAGCACAAATTTGGTCAACGCTCCCTCTCTCGGTGGA 3287
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGATCCCGTGT 1080
Db 3288 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGATCCCGTGT 3347
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTACAG 1140
Db 3348 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTACAG 3407
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGGATTTGTGACCAAGCGAGGCTTGGGCT 1200
Db 3408 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGGATTTGTGACCAAGCGAGGCTTGGGCT 3467
QY 1201 CCTTCTGATGATTTCTATCCAAATTCACAGACAGTGTGGGTGAAGTTGCATCTCTTGGGG 1260
Db 3468 CCTTCTGATGATTTCTATCCAAATTCACAGACAGTGTGGGTGAAGTTGCATCTCTTGGGG 3527
QY 1261 CAGTAAATTCAGCAAGTGTCCGAGCTGCTTCCCAATTG 1301
Db 3528 CAGTAAATTCAGCAAGTGTCCGAGCTGCTTCCCAATTG 3568

RESULT 7
AAD37230
ID AAD37230 standard; DNA; 4182 BP.
XX AAD37230;
AC AAD37230;
XX 21-AUG-2002 (first entry)
DT Human dystrophin minigene delta4173.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.
OS Homo sapiens.
XX WO200183695-A2.
PN 08-NOV-2001.
PD 27-APR-2001; 2001WO-US13677.
PF 28-APR-2000; 2000US-200777P.
PR (XIAO/) XIAO X.
PA XIAO X;
XX XIAO X;
XX WPI; 2002-049342/06.
DR New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a

TZ /

PT New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal

PT New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal

rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -

Example 1; Page 63-65; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence.

Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 CGACTTTCCAGCAGTTCAGAGCAGACGATGATACATAGGCGCTTCAGAGGGAATTGAA 60
2535 CGACTTTCCAGCAGTTCAGAGCAGACGATGATACATAGGCGCTTCAGAGGGAATTGAA 2594
61 AACTAAAGAACCTGTATATCATGTAGTACTCTGTAGAGCTGTACGAATATTTCTGACAGAGCA 120
2595 AACTAAAGAACCTGTATATCATGTAGTACTCTGTAGAGCTGTACGAATATTTCTGACAGAGCA 2654
121 GCCTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTCCCTCTGAGAGAG 180
2655 GCCTTTGGAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTCCCTCTGAGAGAG 2714
181 AGCCAGAGATGTCTACCTGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA 240
2715 AGCCAGAGATGTCTACCTGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA 2774
241 AAAATGAACTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACT 300
2775 AAAATGAACTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCCCTTGAAGACT 2834
301 CCAGGAACTTCAGAGGCCAGGATGAGCTGACCTCAAGCTGCGCAAGCTGAGGTGAT 360
2835 CCAGGAACTTCAGAGGCCAGGATGAGCTGACCTCAAGCTGCGCAAGCTGAGGTGAT 2894
361 CAAGGATCTCTGGCAGCCCGTGGCGGATCTCTCTCATTGACTCTCTCCAGATCACCTCGA 420
2895 CAAGGATCTCTGGCAGCCCGTGGCGGATCTCTCTCATTGACTCTCTCCAGATCACCTCGA 2954
421 GAAAGTCAAGGCACTTCAGAGAGAAATGGCTCTCTGAAAGAGAAAGTGAGCCACGCTCAA 480
2955 GAAAGTCAAGGCACTTCAGAGAGAAATGGCTCTCTGAAAGAGAAAGTGAGCCACGCTCAA 3014
481 TGACCTTGTCTCGCAGCTTACCACCTTGGGCAATTCAGCTCTCACCGCTATAACCTCAGCAC 540
3015 TGACCTTGTCTCGCAGCTTACCACCTTGGGCAATTCAGCTCTCACCGCTATAACCTCAGCAC 3074
541 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCGAGGTGGCCGTGAGAGCCGAGTCTAG 600
3075 TCTGGAAGACCTGAAACACAGATGGAAGCTTCTGCGAGGTGGCCGTGAGAGCCGAGTCTAG 3134
601 GCAGCTGCATGAGCCACAGGAGCTTTGGTTCAGGATCTCTCAGCATCTCTCTTCTTCCACGTC 660
3135 GCAGCTGCATGAGCCACAGGAGCTTTGGTTCAGGATCTCTCAGCATCTCTCTTCTTCCACGTC 3194
661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTTACTATATCAACCA 720
3195 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTTACTATATCAACCA 3254
721 CGAGACTCAAAACAACTTGTCTGGGAGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
```

```
Db 3255 CGAGACTCAAAACAATCTGCTGGGACCATCCCAAAATGACAGAGCTCTTACAGTCTTTAGC 3314
QY 781 TGACCTGAAATATCTCAGATTTCTCAGCTTTAGGACTGGCATGAACTCCGAAAGACTGCA 840
Db 3315 TGACCTGAAATATCTCAGATTTCTCAGCTTTAGGACTGGCATGAACTCCGAAAGACTGCA 3374
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGCGATGTGATGATGCTTTGGACCA 900
Db 3375 GAAGGCCCTTTGCTTGGATCTCTTGGAGCTCTCAGCTGCGATGTGATGATGCTTTGGACCA 3434
QY 901 CAACCTCAGCAAAATGACAGCCCATGGATATCTCGCAGATTTATTAATTTGTTGACCA 960
Db 3435 CAACCTCAGCAAAATGACAGCCCATGGATATCTCGCAGATTTATTAATTTGTTGACCA 3494
QY 961 TATTTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACCTCTCTCTCGGTGGA 1020
Db 3495 TATTTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACCTCTCTCTCGGTGGA 3554
QY 1021 TATGTGCTGAACCTGGCTGTGAATGTTTATGATACGGGAGCAACAGGAGATCCGTGT 1080
Db 3555 TATGTGCTGAACCTGGCTGTGAATGTTTATGATACGGGAGCAACAGGAGATCCGTGT 3614
QY 1081 CCGTCTTTTAAACTGSCATCATTTCCCTGTGTAAGACACATTTGGAAGCAAGTACAG 1140
Db 3615 CCGTCTTTTAAACTGSCATCATTTCCCTGTGTAAGACACATTTGGAAGCAAGTACAG 3674
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGCGCAGGCTGGGCT 1200
Db 3675 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGCGCAGGCTGGGCT 3734
QY 1201 CCGTCTGATGATTTCTATCCAAATTCGAAGACAGTGGGTGAAGTTCATCCTTTGGGG 1260
Db 3735 CCGTCTGATGATTTCTATCCAAATTCGAAGACAGTGGGTGAAGTTCATCCTTTGGGG 3794
QY 1261 CAGTAACTTCAGCAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
Db 3795 CAGTAACTTCAGCAAGTGTCCGGAGCTGCTTCCAAATTTG 3835
```

RESULT 9

AAAD37258
ID AAD37258 standard; DNA; 4498 BP.

AC AAD37258;

XX 21-AUG-2002 (first entry)

DE Adeno-associated virus vector plasmid, AAV-MCK-3531.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DVD; BMD;

KW Becker muscular dystrophy; ds.

OS Chimeric - Homo sapiens.

OS Chimeric - Unidentified.

PN WO200183695-A2.

PD 08-NOV-2001.

PF 27-APR-2001; 2001WO-US13677.

PR 28-APR-2000; 2000US-200777P.

PA (XIAO/) XIAO X.

PI Xiao X;

XX WPI; 2002-049342/06.

PT New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain.

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
 PT dystrophin gene

XX Example 1; Page 62-63; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified
 CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid construct containing human dystrophin minigenes, a muscle
 CC creatine kinase (MCK) promoter and a small polyA signal sequence.

XX Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 4498;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCAGCAGTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 60
 DB 2557 CGACTTTCAGCAGTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 2616
 QY 61 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTCACAGAGCA 120
 DB 2617 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTCACAGAGCA 2676
 QY 121 GCCTTTGGAGAGCTAGAGAACTCTACAGAGAGCCAGAGAGCTCCCTCTGAGAGAG 180
 DB 2677 GCCTTTGGAGAGCTAGAGAACTCTACAGAGAGCCAGAGAGCTCCCTCTGAGAGAG 2736
 QY 181 AGCCAGAAATGTCACCTGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
 DB 2737 AGCCAGAAATGTCACCTGGCTTCTACGAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 2796
 QY 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 300
 DB 2797 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 2856
 QY 301 CCAGGAATCTCAGAGAGCCAGAGTGAAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 360
 DB 2857 CCAGGAATCTCAGAGAGCCAGAGTGAAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 2916
 QY 361 CAAGGGATCTGGCAGAGCCGCTGGCGGATCTCTCATTTGACTCTCTCCAAAGATCACCTCGA 420
 DB 2917 CAAGGGATCTGGCAGAGCCGCTGGCGGATCTCTCATTTGACTCTCTCCAAAGATCACCTCGA 2976
 QY 421 GAAAGTCAAGGCACTTCGAGAGAGAAATTTGGCCTCTGAAAGAGAACTGAGGCCACGTCAA 480
 DB 2977 GAAAGTCAAGGCACTTCGAGAGAGAAATTTGGCCTCTGAAAGAGAACTGAGGCCACGTCAA 3036
 QY 481 TGACCTTGCTCGCAGCTTACCACTTTGGCAGTTCAGCTCTCACCGTATAAAGCTCAGAC 540
 DB 3037 TGACCTTGCTCGCAGCTTACCACTTTGGCAGTTCAGCTCTCACCGTATAAAGCTCAGAC 3096
 QY 541 TCTGGAAGACTGAACACCAAGATGGAAGTCTTGCAGGTGGCGCTCGAGGACCGAGTCAAG 600
 DB 3097 TCTGGAAGACTGAACACCAAGATGGAAGTCTTGCAGGTGGCGCTCGAGGACCGAGTCAAG 3156
 QY 601 CGACTGATGAAGCCACAGGAGCTTTGGTCCAGGATCTCAGACCTTTCTTTCCACGTC 660
 DB 3157 CGACTGATGAAGCCACAGGAGCTTTGGTCCAGGATCTCAGACCTTTCTTTCCACGTC 3216
 QY 661 TGTCCAGGCTCCCTGGGAGAGAGCACTCTCGCCAAACAAAGTGCCTCTATATCAACCA 720
 DB 3217 TGTCCAGGCTCCCTGGGAGAGAGCACTCTCGCCAAACAAAGTGCCTCTATATCAACCA 3276
 QY 721 CGAGACTCAAAACAACTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGCTTTAGC 780

DB 3277 CGAGACTCAAAACAACTGTCTGGAGCCATCCCAAAATGACAGAGCTCTACCAAGCTTTAGC 3336
 QY 781 TGACCTGAAATATGTCTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCA 840
 DB 3337 TGACCTGAAATATGTCTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCA 3396
 QY 841 GAAGGCCCTTTGCTTGGATCTCTTTGAGCCCTGTGAGCTGTGATGCTTGGACACAGA 900
 DB 3397 GAAGGCCCTTTGCTTGGATCTCTTTGAGCCCTGTGAGCTGTGATGCTTGGACACAGA 3456
 QY 901 CAACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCAAC 960
 DB 3457 CAACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCAAC 3516
 QY 961 TATTTATGACCGCTCGAGCAGAGCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 1020
 DB 3517 TATTTATGACCGCTCGAGCAGAGCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 3576
 QY 1021 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGTGT 1080
 DB 3577 TATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAGGAGGATCCGTGT 3636
 QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACAATTTGGAAGACAAGTACAG 1140
 DB 3637 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACAATTTGGAAGACAAGTACAG 3696
 QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACCGAGCGAGCTGGGCT 1200
 DB 3697 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACCGAGCGAGCTGGGCT 3756
 QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCCAGACAGTTTGGGTGAAGTTGCAATCTTTGGGG 1260
 DB 3757 CCTTCTGCATGATTTCTATCCAAATTTCCAGACAGTTTGGGTGAAGTTGCAATCTTTGGGG 3816
 QY 1261 CAGTAACATTTGAGCAGAGTGTCCGAGCTGCTTCCAAATTTG 1301
 DB 3817 CAGTAACATTTGAGCAGAGTGTCCGAGCTGCTTCCAAATTTG 3857

RESULT 10

AAD37257

ID AAD37257 standard; DNA; 4825 BP.

XX AAD37257;

XX 21-AUG-2002 (first entry)

XX Adeno-associated virus vector plasmid, AAV-MCK-delta3849.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX Chimeric - Homo sapiens.

OS Chimeric - Unidentified.

PN WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

PI WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
 PT dystrophin gene -
 XX
 XX
 PS
 XX
 XX
 CC The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified
 CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid construct containing human dystrophin minigenes, a muscle
 CC creatine kinase (MCK) promoter and a small polyA signal sequence.
 XX
 XX
 SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 4825;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTCCAGCAGTTCAAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 60
 DB 2884 CGACTTCCAGCAGTTCAAGAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 2843

QY 61 AACTAAAGAACTGPAATCATGAGTACTCTTGAAGACTGTACGAATATTTCTGACAGACGA 120
 DB 2944 AACTAAAGAACTGPAATCATGAGTACTCTTGAAGACTGTACGAATATTTCTGACAGACGA 3003

QY 121 GCCTTTGAAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTGAGGAGAG 180
 DB 3004 GCCTTTGAAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTGAGGAGAG 3063

QY 181 AGCCAGAAATGCTCACTCGGCTTTACGAAGACAGGCTGAGAGGCTCAATACTGAGTGGGA 240
 DB 3064 AGCCAGAAATGCTCACTCGGCTTTACGAAGACAGGCTGAGAGGCTCAATACTGAGTGGGA 3123

QY 241 AAAATGAACTGCACTCGCTGACTGCGCAGAGAAAATAGATGAGACCTTTGAAAGACT 300
 DB 3124 AAAATGAACTGCACTCGCTGACTGCGCAGAGAAAATAGATGAGACCTTTGAAAGACT 3183

QY 301 CCAGGAATCTCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGGCCAAAGCTGAGGTGAT 360
 DB 3184 CCAGGAATCTCAAGAGGCCAGGATGAGCTGAGCTCAAGCTGGCCAAAGCTGAGGTGAT 3243

QY 361 CAAGGATCTCGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 420
 DB 3244 CAAGGATCTCGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACCTCGA 3303

QY 421 GAAAGTCAGGCACCTTCAGAGAGAAATTCGCTCTGAAGAGAACCTGAGCCACGTCAA 480
 DB 3304 GAAAGTCAGGCACCTTCAGAGAGAAATTCGCTCTGAAGAGAACCTGAGCCACGTCAA 3363

QY 481 TGACCTTGTCTGCCAGCTTACCATTCTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC 540
 DB 3364 TGACCTTGTCTGCCAGCTTACCATTCTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC 3423

QY 541 TCTGGAACACCTTGAACACAGATGGAAGCTTCTGAGGTGGCCGTCAGGACCGAGTCAG 600
 DB 3424 TCTGGAACACCTTGAACACAGATGGAAGCTTCTGAGGTGGCCGTCAGGACCGAGTCAG 3483

QY 601 GCAGCTGATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGACATTTCTTTCCAGCTC 660
 DB 3484 GCAGCTGATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGACATTTCTTTCCAGCTC 3543

QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTATATCAACCA 720
 DB 3544 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTTACTATATCAACCA 3603

QY 721 CGAGACTCAACAACCTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780

DB 3604 CGAGACTCAACAACCTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3663
 QY 781 TGACCTGAATAATGTCCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGCA 840
 DB 3664 TGACCTGAATAATGTCCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGCA 3723
 QY 841 GAAGGCCCTTTCTCTGGATCTCTTGAAGCTGTCTCAGCTGATGTGATGCTTGGACAGCA 900
 DB 3724 GAAGGCCCTTTCTCTGGATCTCTTGAAGCTGTCTCAGCTGATGTGATGCTTGGACAGCA 3783
 QY 901 CAACCTCAAGCAAAATGACCCACCCATGATATCTGAGATTTAATTTGTTGACCAAC 960
 DB 3784 CAACCTCAAGCAAAATGACCCACCCATGATATCTGAGATTTAATTTGTTGACCAAC 3843
 QY 961 TATTTATGACCCCTGGAGAGACCAACAATTTGGTCAACGTCCTCTCTGGGTGGA 1020
 DB 3844 TATTTATGACCCCTGGAGAGACCAACAATTTGGTCAACGTCCTCTCTGGGTGGA 3903
 QY 1021 TATGTGCTGAACCTGGCTGCTCAATTTGTTATGATACGGAGCAACAGGGAGATCCGTGT 1080
 DB 3904 TATGTGCTGAACCTGGCTGCTCAATTTGTTATGATACGGAGCAACAGGGAGATCCGTGT 3963
 QY 1081 CTTGTCTTTTAAATCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAACTACAG 1140
 DB 3964 CTTGTCTTTTAAATCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAACTACAG 4023
 QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCGAGGCTGGGCT 1200
 DB 4024 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCGAGGCTGGGCT 4083
 QY 1201 CTTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGTGAAGTTGCATCTTTTGGGG 1260
 DB 4084 CTTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGTGAAGTTGCATCTTTTGGGG 4143
 QY 1261 CAGTAACATTTGACCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
 DB 4144 CAGTAACATTTGACCAAGTGTCCGAGCTGCTTCCAAATTTG 4184

RESULT 11
 AAD37263
 ID AAD37263 standard; DNA; 4848 BP.
 XX
 AC AAD37263;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
 XX
 KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX
 OS Chimeric - Homo sapiens.
 OS Chimeric - Cytomegalovirus.
 OS Chimeric - Unidentified.
 XX
 PN WO200183695-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US13677.
 XX
 PR 28-APR-2000; 2000US-200777P.
 XX
 PA (X180/) XIAO X.
 XX
 PI Xiao X;
 XX
 DR WPI; 2002-049342/06.
 XX
 PT New dystrophin minigene for treating Duchenne or Becker muscular

dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

Example 1; Page 68-70; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a cytomegalovirus (CMV) promoter and a small polyA signal sequence.

Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 4848;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAGAGGGAATTGAA 60
 2907 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAGAGGGAATTGAA 2966

61 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTCAGAGAGCA 120
 2967 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTCAGAGAGCA 3026

121 GCCTTTGGAAGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCCCTCTGAGAGAG 180
 3027 GCCTTTGGAAGACTAGAGAACTTACAGAGAGCCAGAGAGCTGCCCTCTGAGAGAG 3086

181 AGCCAGAAATGTCACTCGGCTTCTCGAAGCAGGCTGAGGAGTCAATCTAGTGGGA 240
 3087 AGCCAGAAATGTCACTCGGCTTCTCGAAGCAGGCTGAGGAGTCAATCTAGTGGGA 3146

241 AAAATTGAACTCGCTCACTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 300
 3147 AAAATTGAACTCGCTCACTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 3206

301 CAGGAACTTCAAGAGCCACGAGTACGAGTGGACCTCAAGCTGCGCCAGCTGAGGTGAT 360
 3207 CAGGAACTTCAAGAGCCACGAGTACGAGTGGACCTCAAGCTGCGCCAGCTGAGGTGAT 3266

361 CAAGGATCTGCGAGCCGCTGGCGATCTCTCTCAATGACTCTCTCAAGATCACTCGA 420
 3267 CAAGGATCTGCGAGCCGCTGGCGATCTCTCTCAATGACTCTCTCAAGATCACTCGA 3326

421 GAAAGTCAAGGCACTTGGAGGAAATTTGGCCCTCTGAAAGAACGCTGAGCCAGCTCAA 480
 3327 GAAAGTCAAGGCACTTGGAGGAAATTTGGCCCTCTGAAAGAACGCTGAGCCAGCTCAA 3386

481 TGACCTTGCTCGCAGCTTACCACTTTGGCACTTCACTGCTCACTGCTCACTGCTCACTG 540
 3387 TGACCTTGCTCGCAGCTTACCACTTTGGCACTTCACTGCTCACTGCTCACTGCTCACTG 3446

541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTTGCGAGGTGGCGTCGAGGACCGAGTCAG 600
 3447 TCTGGAAGACCTGAACACCAAGATGGAAGCTTTGCGAGGTGGCGTCGAGGACCGAGTCAG 3506

601 GCAGCTCATGAGCCACAGGACTTTGGTCCAGCATCTCAGCATCTCTCTTCTTCCACGTC 660
 3507 GCAGCTCATGAGCCACAGGACTTTGGTCCAGCATCTCAGCATCTCTCTTCTTCCACGTC 3566

661 TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACCA 720
 3567 TGTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATATCAACCA 3626

QY 721 CGAGACTCAAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
 DB 3627 CGAGACTCAAAACAACTTGCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 3686

QY 781 TGACCTGAATATGTGATGATCTCTGAGTCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
 DB 3687 TGACCTGAATATGTGATGATCTCTGAGTCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3746

QY 841 GAAGGCCCTTTGCTTGATCTCTTGAGCCCTGTGAGCTGATGTGATGCTTGGACCCAGCA 900
 DB 3747 GAAGGCCCTTTGCTTGATCTCTTGAGCCCTGTGAGCTGATGTGATGCTTGGACCCAGCA 3806

QY 901 CAACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCCAC 960
 DB 3807 CAACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTTGTTGACCCAC 3866

QY 961 TATTTATGACCGCTGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 1020
 DB 3867 TATTTATGACCGCTGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 3926

QY 1021 TATGTGCTGAATGCTGCTGATGTTTATGATACGGGACGACAGGAGGATCCGTGT 1080
 DB 3927 TATGTGCTGAATGCTGCTGATGTTTATGATACGGGACGACAGGAGGATCCGTGT 3986

QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAATTTCGAAAGACAAATGACAG 1140
 DB 3987 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACAATTTCGAAAGACAAATGACAG 4046

QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACGAGCGAGGCTGGGCT 1200
 DB 4047 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACGAGCGAGGCTGGGCT 4106

QY 1201 CCTTCTGATGATTTCTATCCAAATTCAGACAGTTGGGTGAAGTTGATCTCTTTGGGG 1260
 DB 4107 CCTTCTGATGATTTCTATCCAAATTCAGACAGTTGGGTGAAGTTGATCTCTTTGGGG 4166

QY 1261 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAAATTG 1301
 DB 4167 CAGTAACATTGAGCCCAAGTGTCCGGAGCTGCTTCCAAATTG 4207

RESULT 12
 AAD37256
 ID AAD37256 standard; DNA; 4966 BP.
 XX
 AC AAD37256;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE Adeno-associated virus vector plasmid, AAV-MCK-delta3990.
 XX
 KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX
 OS Chimeric - Homo sapiens.
 OS Chimeric - Unidentified.
 XX
 PN WO200183695-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US13677.
 XX
 PR 28-APR-2000; 2000US-20077P.
 XX
 PA (XIAO)/ XIAO X.
 XX
 PI Xiao X;
 XX
 DR WPI; 2002-049342/06.
 XX
 PT New dystrophin minigene for treating Duchenne or Becker muscular

dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -

Example 1; Page 59-60; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence.

Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;

Query Match 100.0%; Score 1301; DB 24; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACTTTCCAGCAGTTTCAGAGCAGACGATGACATAGGCGCTTCAGAGGGAATTGAA 60
3025 CGACTTTCCAGCAGTTTCAGAGCAGACGATGACATAGGCGCTTCAGAGGGAATTGAA 3084

61 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGAGCTGTACGAATATTTCTGACAGAGCA 120
3085 AACTAAAGAACCTGTAATCATGAGTACTCTTGAGAGCTGTACGAATATTTCTGACAGAGCA 3144

121 GCCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAG 180
3145 GCCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAG 3204

181 AGCCAGAGTCTCACTGGCTTTCTACAAAGAGCGGTGAGAGGTCAATACTCAGTGGGA 240
3205 AGCCAGAGTCTCACTGGCTTTCTACAAAGAGCGGTGAGAGGTCAATACTCAGTGGGA 3264

241 AAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 300
3265 AAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 3324

301 CCAGGAACTTCAAGAGCCAGGATGAGTGGACCTCAAGCTGGCCAGCTGAGTGAT 360
3325 CCAGGAACTTCAAGAGCCAGGATGAGTGGACCTCAAGCTGGCCAGCTGAGTGAT 3384

361 CAAGGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTGACTCTCTCAAGATCACCTCGA 420
3385 CAAGGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTGACTCTCTCAAGATCACCTCGA 3444

421 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTGAAAGAGAACTGAGCAGCAGTCAA 480
3445 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTGAAAGAGAACTGAGCAGCAGTCAA 3504

481 TGACCTTGTCTCGCAGCTTACCACTTTGGGCACTTACGCTCTCAAGGATTAACCTCAGCAC 540
3505 TGACCTTGTCTCGCAGCTTACCACTTTGGGCACTTACGCTCTCAAGGATTAACCTCAGCAC 3564

541 TCTGGAAGACTGAAACCAAGATGGAAGCTTTCTGAGGTGGCGCTCGAGGACCGAGTCAAG 600
3565 TCTGGAAGACTGAAACCAAGATGGAAGCTTTCTGAGGTGGCGCTCGAGGACCGAGTCAAG 3624

601 GCAGCTGCATGAGCCACAGGAGCTTTGGTCCAGGATCTCAGCAGCTTTCTTTCCACGTC 660
3625 GCAGCTGCATGAGCCACAGGAGCTTTGGTCCAGGATCTCAGCAGCTTTCTTTCCACGTC 3684

661 TGTCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720
3685 TGTCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 3744

QY 721 CGAGACTCAAAACAATCTGCTGGGACCAATCCCAAATGACAGAGCTCTACCAAGTCTTTAGC 780
DB 3745 CGAGACTCAAAACAATCTGCTGGGACCAATCCCAAATGACAGAGCTCTACCAAGTCTTTAGC 3804
QY 781 TGACCTGAAATGTCAGATTTCTCAGCTTATAGACTGCCATGAACTCCGAAAGCTGCA 840
DB 3805 TGACCTGAAATGTCAGATTTCTCAGCTTATAGACTGCCATGAACTCCGAAAGCTGCA 3864
QY 841 GAAGGCCCTTTGGCTTGGATCTCTTGGAGCTGTAGCTGCGATGTGATGCTTTGGACCA 900
DB 3865 GAAGGCCCTTTGGCTTGGATCTCTTGGAGCTGTAGCTGCGATGTGATGCTTTGGACCA 3924
QY 901 CAACCTCAAGAAATGACAGCCCATGGATATCTCGAGATTAATTAATTTGACCAAC 960
DB 3925 CAACCTCAAGAAATGACAGCCCATGGATATCTCGAGATTAATTAATTTGACCAAC 3984
QY 961 TATTTATCAGCCCTGGAGCAAGAGCACAACAATTTGGTCAACGCTCCTCTCTCGGTGGA 1020
DB 3985 TATTTATCAGCCCTGGAGCAAGAGCACAACAATTTGGTCAACGCTCCTCTCTCGGTGGA 4044
QY 1021 TATGTCCTGAATGGCTGCTGAAATGTTTATGATACGGAGCAACAGGAGGATCCGTGT 1080
DB 4045 TATGTCCTGAATGGCTGCTGAAATGTTTATGATACGGAGCAACAGGAGGATCCGTGT 4104
QY 1081 CCGTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGCAAGTACAG 1140
DB 4105 CCGTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGACACATTTGGAAGCAAGTACAG 4164
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 1200
DB 4165 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCT 4224
QY 1201 CCTTCTGATGATTTATCCAAATTCGAAGACAGTGGGTGAGTTGATTCATCTTTGGGG 1260
DB 4225 CCTTCTGATGATTTATCCAAATTCGAAGACAGTGGGTGAGTTGATTCATCTTTGGGG 4284
QY 1261 CAGTAACATTTGAGCCCAAGTGTCCGAGAGCTGCTTCCAATTTG 1301
DB 4285 CAGTAACATTTGAGCCCAAGTGTCCGAGAGCTGCTTCCAATTTG 4325

RESULT 13

AAD37262

ID AAD37262 standard; DNA; 4990 BP.

XX AC AAD37262;

XX XX

DT 21-AUG-2002 (first entry)

XX XX

DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3990.

XX XX

KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX XX

OS Chimeric - Homo sapiens.

OS Chimeric - Cytomegalovirus.

OS Chimeric - Unidentified.

XX XX

PN WO200183695-A2.

XX XX

PD 08-NOV-2001.

XX XX

PF 27-APR-2001; 2001WO-US13677.

XX XX

PR 28-APR-2000; 2000US-200777P.

XX XX

PA (XIAO/) XIAO X.

XX XX

PI Xiao X;

XX XX

DR WPI; 2002-049342/06.

XX XX

PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
XX Example 1; Page 67-68; 71pp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence.
XX
XX Sequence 4990 BP; 1439 A; 1185 C; 1208 G; 1158 T; 0 other;
SQ
Query Match 100.0%; Score 1301; DB 24; Length 4990;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGATACATAGGCGCTTCAAGAGGGAATTGAA 60
DB 3049 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGATACATAGGCGCTTCAAGAGGGAATTGAA 3108
QY 61 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTCTACGAATATTTCTGACAGAGCA 120
DB 3109 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTCTACGAATATTTCTGACAGAGCA 3168
QY 121 GCCTTTGGAGAGTACTAGAGAACTCTTACAGAGAGCCAGAGAGCTGCCTCTCTGAGGAGAG 180
DB 3169 GCCTTTGGAGAGTACTAGAGAACTCTTACAGAGAGCCAGAGAGCTGCCTCTCTGAGGAGAG 3228
QY 181 AGCCAGATGTCACCTCGGCTTCTAGAAAGCAGAGCTGAGGAGGTCAATGAGTGGGA 240
DB 3229 AGCCAGATGTCACCTCGGCTTCTAGAAAGCAGAGCTGAGGAGGTCAATGAGTGGGA 3288
QY 241 AAAATTGAACCTGACTCGGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 300
DB 3289 AAAATTGAACCTGACTCGGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 3348
QY 301 CCAGGAACCTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCCAGCTGAGGTGAT 360
DB 3349 CCAGGAACCTCAAGAGGCCACCGATGAGTGGACCTCAAGCTGCGCCAGCTGAGGTGAT 3408
QY 361 CAAGGATCTTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAGAGTCACTCGA 420
DB 3409 CAAGGATCTTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAGAGTCACTCGA 3468
QY 421 GAAAGTCAAGGCATCTCGAGGAGAAATTCGGCTCTGAAAGAGAAACGTGAGCCACGTCAA 480
DB 3469 GAAAGTCAAGGCATCTCGAGGAGAAATTCGGCTCTGAAAGAGAAACGTGAGCCACGTCAA 3528
QY 481 TGACCTTGTCTGCGCAGTTACCACTTTGGCATTGAGCTCTACCGTATAACCTCAGCAC 540
DB 3529 TGACCTTGTCTGCGCAGTTACCACTTTGGCATTGAGCTCTACCGTATAACCTCAGCAC 3588
QY 541 TCTGGAAGACCTGAAACACAGATGGAAGCTTTCTGCAAGTGGCGCTCGAGGACCGAGTCCAG 600
DB 3589 TCTGGAAGACCTGAAACACAGATGGAAGCTTTCTGCAAGTGGCGCTCGAGGACCGAGTCCAG 3648
QY 601 GCAGCTGCATGAAGCCCAAGGAGCTTTGGTTCAGCATCTCAGACATTTCTTTCCACGTC 660
DB 3649 GCAGCTGCATGAAGCCCAAGGAGCTTTGGTTCAGCATCTCAGACATTTCTTTCCACGTC 3708
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCATCTCTCGCAAAACAAAGTCCCTACTATATCAACCA 720
DB 3709 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCTGCCAAACAAAGTCCCTACTATATCAACCA 3768

QY 721 CGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 780
DB 3769 CGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGC 3828
QY 781 TGACCTGAATAATGTCCAGATTCCTAGCTTATAGGACTGCCATGAAGAACTCCGAGAGACTGCA 840
DB 3829 TGACCTGAATAATGTCCAGATTCCTAGCTTATAGGACTGCCATGAAGAACTCCGAGAGACTGCA 3888
QY 841 GAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGACAGCA 900
DB 3889 GAAGGCCCTTTGCTGGATCTCTTGAGCCTGTGAGCTGATGATGCTTGGACAGCA 3948
QY 901 CAACTCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATATTAATTTGTTGACCAAC 960
DB 3949 CAACTCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATATTAATTTGTTGACCAAC 4008
QY 961 TATTTATGACCGCTGGAGCAGAGACAAACAATTTGGTCAACGCTCCCTCTCTGCGTGA 1020
DB 4009 TATTTATGACCGCTGGAGCAGAGACAAACAATTTGGTCAACGCTCCCTCTCTGCGTGA 4068
QY 1021 TATGTGTCTGAAGTGGCTGCTGAATGTTTATGATCGGACGAAACAGGAGGATCCGTGT 1080
DB 4069 TATGTGTCTGAAGTGGCTGCTGAATGTTTATGATCGGACGAAACAGGAGGATCCGTGT 4128
QY 1081 CTTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGAGCAAGTACAG 1140
DB 4129 CTTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGAGCAAGTACAG 4188
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTACCCAGCGCAGGCTGGGCT 1200
DB 4189 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTACCCAGCGCAGGCTGGGCT 4248
QY 1201 CTTTCTGATGATTCATCCAAATTCGAAGACAGTGGTGAAGTGCATCTCTTTGGGG 1260
DB 4249 CTTTCTGATGATTCATCCAAATTCGAAGACAGTGGTGAAGTGCATCTCTTTGGGG 4308
QY 1261 CAGTAACATTTGACCCAGTGTCCGAGCTGCTTCCAAATTTG 1301
DB 4309 CAGTAACATTTGACCCAGTGTCCGAGCTGCTTCCAAATTTG 4349
RESULT 14
AAD37264
ID AAD37264 standard; DNA; 5060 BP.
XX AAD37264;
AC AAD37264;
DT 21-AUG-2002 (first entry)
XX Adeno-associated virus (AAV) vector plasmid, AAV-E-CMV-3849.
DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
XX Chimeric - Unidentified.
XX WO200183695-A2.
XX 08-NOV-2001.
XX 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
XX Xiao X;
XX WPI; 2002-049342/06.

XX PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene
XX
PS Example 1; Page 70-71; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domain, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter
CC and a small polyA signal sequence.
XX
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;
Query Match 100.0%; Score 1301; DB 24; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2Y 1 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 60
2b 3119 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 3178
2Y 61 AACTAAGAACCTGTATCATGTAGTACTCTTGAGACTGTACGAATATTTCTCACAGAGCA 120
2b 3179 AACTAAGAACCTGTATCATGTAGTACTCTTGAGACTGTACGAATATTTCTCACAGAGCA 3238
2Y 121 GCCTTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTCCAGAGAG 180
2b 3239 GCCTTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTCCAGAGAG 3298
2Y 181 AGCCAGAGATGTCACTCGGCTTACGAAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 240
2b 3299 AGCCAGAGATGTCACTCGGCTTACGAAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 3358
2Y 241 AAAATTGAACCTCGACTCGCTGACCTGAGAGAGAAATAGATGAGACCTTTGAAAGACT 300
2b 3359 AAAATTGAACCTCGACTCGCTGACCTGAGAGAGAAATAGATGAGACCTTTGAAAGACT 3418
2Y 301 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGAACCTTCAAGCTTCGCCAAGCTGAGGTGAT 360
2b 3419 CCAGGAACCTTCAAGAGGCCACCGATGAGCTGGAACCTTCAAGCTTCGCCAAGCTGAGGTGAT 3478
2Y 361 CAAGGATCTGCGACCGCTGGCGCATCTCTCATTTGACTCTCTCCAGATCACTCGA 420
2b 3479 CAAGGATCTGCGACCGCTGGCGCATCTCTCATTTGACTCTCTCCAGATCACTCGA 3538
2Y 421 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAACGCTGAGCAGCTCAA 480
2b 3539 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAACGCTGAGCAGCTCAA 3598
2Y 481 TGACCTTGTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC 540
2b 3599 TGACCTTGTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC 3658
2Y 541 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 600
2b 3659 TCTGGAAGACCTGAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 3718
2Y 601 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGACTTTCTTTCCACGTC 660
2b 3719 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGACTTTCTTTCCACGTC 3778
2Y 661 TGTCGAGGTCCTTCGGGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720

Db 3779 TGTCCAGGTCCTTCGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 3838
Qy 721 CGAGACTCAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAC 780
Db 3839 CGAGACTCAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTTTTAC 3898
Qy 781 TGACCTGAATATATGCTCAGATTCTCAGCTTATAGAGTGCATGAACTCCGAAGACTGCA 840
Db 3899 TGACCTGAATATATGCTCAGATTCTCAGCTTATAGAGTGCATGAACTCCGAAGACTGCA 3958
Qy 841 GAAGGCTTGTCTGGATCTCTTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 900
Db 3959 GAAGGCTTGTCTGGATCTCTTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 4018
Qy 901 CAACCTCAAGCAAAATGACAGGCCCATGATGATATCCCTGCAGATTATTAATTGTTGACCAC 960
Db 4019 CAACCTCAAGCAAAATGACAGGCCCATGATATCCCTGCAGATTATTAATTGTTGACCAC 4078
Qy 961 TATTATGACGGCTGGAGCAAGAGCAACAATTTGGTCAAGTCCCTCTCTGCGTGA 1020
Db 4079 TATTATGACGGCTGGAGCAAGAGCAACAATTTGGTCAAGTCCCTCTCTGCGTGA 4138
Qy 1021 TATGTGTCTGAATCTGCTGCTGAATGTTATGATACGGGACGAAACAGGAGGATCCGCTGT 1080
Db 4139 TATGTGTCTGAATCTGCTGCTGAATGTTATGATACGGGACGAAACAGGAGGATCCGCTGT 4198
Qy 1081 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 1140
Db 4199 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAGTACAG 4258
Qy 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGAGCGAGCTGGGCT 1200
Db 4259 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCGAGCGAGCTGGGCT 4318
Qy 1201 CCTTCTGCATGATTCTATCCAAATTTCCAGACAGTTGGGTGAAAGTTGCATCTTTGGGG 1260
Db 4319 CCTTCTGCATGATTCTATCCAAATTTCCAGACAGTTGGGTGAAAGTTGCATCTTTGGGG 4378
Qy 1261 CAGTAAACATTGAGCAAGTGTCCGGAGCTGCTTCCAAATTG 1301
Db 4379 CAGTAAACATTGAGCAAGTGTCCGGAGCTGCTTCCAAATTG 4419
RESULT 15
AAD37255
ID AAD37255 standard; DNA; 5149 BP.
XX
AC AAD37255;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta4173.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
FN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 23:51:42 ; Search time 425.945 Seconds
(without alignments)
11131.886 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100

Perfect score: 1301

Sequence: 1 cgcatttcgcagcttcaga.....ccggagctgctcccaatttg 1301

Scoring table: IDENTITY NUC

Gap 10.0, Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1301	100.0	1821	13	US-09-845-416-13
2	1301	100.0	2169	13	US-09-845-416-4
3	1301	100.0	3510	13	US-09-845-416-12
4	1301	100.0	3531	13	US-09-845-416-10
5	1301	100.0	3858	13	US-09-845-416-9
6	1301	100.0	3999	13	US-09-845-416-6
7	1301	100.0	4182	13	US-09-845-416-2
8	1301	100.0	4476	13	US-09-845-416-31
9	1301	100.0	4498	13	US-09-845-416-30
10	1301	100.0	4825	13	US-09-845-416-29
11	1301	100.0	4848	13	US-09-845-416-35
12	1301	100.0	4966	13	US-09-845-416-28
13	1301	100.0	4990	13	US-09-845-416-34
14	1301	100.0	5060	13	US-09-845-416-36
15	1301	100.0	5149	13	US-09-845-416-27

16	1301	100.0	5339	13	US-10-149-736-40	Sequence 40, Appl
17	1301	100.0	5462	13	US-10-149-736-41	Sequence 41, Appl
18	1301	100.0	8689	13	US-10-149-736-42	Sequence 42, Appl
19	1301	100.0	11058	13	US-09-845-416-1	Sequence 1, Appl
20	1301	100.0	11443	13	US-10-149-736-44	Sequence 44, Appl
21	1301	100.0	12057	13	US-10-149-736-47	Sequence 47, Appl
22	1301	100.0	13957	10	US-09-782-378A-22	Sequence 22, Appl
23	1301	100.0	13957	10	US-09-880-107-2284	Sequence 2284, Ap
24	1301	100.0	13957	13	US-10-149-736-1	Sequence 1, Appl
25	1301	100.0	14082	13	US-10-341-434-108	Sequence 108, App
26	1135.2	87.3	13815	13	US-10-149-736-2	Sequence 2, Appl
27	1017.6	78.2	3446	13	US-09-845-416-14	Sequence 14, Appl
28	1017.6	78.2	4414	13	US-09-845-416-32	Sequence 32, Appl
29	1016.6	78.1	5417	13	US-10-149-736-39	Sequence 39, Appl
30	1016	78.1	1434	13	US-09-845-416-15	Sequence 15, Appl
31	567.2	43.6	11096	13	US-10-149-736-4	Sequence 4, Appl
32	553.8	42.6	10302	10	US-09-782-378A-23	Sequence 23, Appl
33	553.8	42.6	10302	13	US-10-149-736-3	Sequence 3, Appl
34	545	41.9	16531	13	US-10-101-510-667	Sequence 667, App
35	472	36.3	887	13	US-10-149-736-35	Sequence 35, Appl
36	324	24.9	324	13	US-10-149-736-33	Sequence 33, Appl
37	289	22.2	387	13	US-10-149-736-32	Sequence 32, Appl
38	216	16.6	216	13	US-10-149-736-34	Sequence 34, Appl
39	114	8.8	114	13	US-10-149-736-45	Sequence 45, Appl
C 40	102.8	7.9	275	13	US-10-029-386-16813	Sequence 16813, A
C 41	102.8	7.9	587	13	US-10-029-386-3113	Sequence 3113, Ap
42	54.2	4.2	449	11	US-09-518-995-24084	Sequence 24084, A
43	54.2	4.2	2200	12	US-10-108-260A-194	Sequence 194, App
44	54.2	4.2	2247	10	US-09-960-253-157	Sequence 157, App
45	42.8	3.3	1690	13	US-10-117-722-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1

US-09-845-416-13
; Sequence 13, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-13

Query Match	100.0%	Score 1301;	DB 13;	Length 1821;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1301;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	CGACATTCACAGCTTCAGACGACGACGATGATACATAGGCGCTTCAAGAGGGAAATTGAA	60	
DB	103	CGACATTCACAGCTTCAGACGACGACGATGATACATAGGCGCTTCAAGAGGGAAATTGAA	162	
QY	61	AACTAAAGAACTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGCAGAGCA	120	
DB	163	AACTAAAGAACTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGCAGAGCA	222	
QY	121	GCCTTTGAAGACTAGAGAACTCTACAGAGCCCGACAGAGCTGCTCTCTGAGGAGAG	180	
DB	223	GCCTTTGAAGACTAGAGAACTCTACAGAGCCCGACAGAGCTGCTCTCTGAGGAGAG	282	
QY	181	AGCCCGAGAACTGCTACTCGGCTTCTACGAAAGAGGCTGAGGAGTCAATCTGAGTGGGA	240	

```
Db 283 AGCCGAGAAATGTCACCTGGCTTCTACGAAGCAGGCTGAGGAGTCAATATCTGAGTGGGA 342
QY 241 AAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAGACT 300
Db 343 AAAATTGAACCTGCACTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAGACT 402
QY 301 CGAGGAACCTTCAAGAGCCACCGATGAGCTGAGCTCAAGCTGCGCCCAAGCTGAGGTGAT 360
Db 403 CGAGGAACCTTCAAGAGCCACCGATGAGCTGAGCTCAAGCTGCGCCCAAGCTGAGGTGAT 462
QY 361 CAAGGGATCTGCGACCGCTGGCGATCTCTCTCATTTGACTCTCTCCAAGATCACTCGA 420
Db 463 CAAGGGATCTGCGACCGCTGGCGATCTCTCTCATTTGACTCTCTCCAAGATCACTCGA 522
QY 421 GAAAGTCAGGACCTTCGAGGAGAAATGGCGCTCTGAAAGAGACGTGAGCCAGTCAA 480
Db 523 GAAAGTCAGGACCTTCGAGGAGAAATGGCGCTCTGAAAGAGACGTGAGCCAGTCAA 582
QY 481 TGACCTTGCTCGCCAGCTTACCACTTTGGCATTGAGCTCTCACCGTATAAAGCTCAGCAC 540
Db 583 TGACCTTGCTCGCCAGCTTACCACTTTGGCATTGAGCTCTCACCGTATAAAGCTCAGCAC 642
QY 541 TCTGGAAGACCTTGAACACCGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 600
Db 643 TCTGGAAGACCTTGAACACCGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 702
QY 601 GCAGCTGTCATGAAAGCCACAGGACCTTTGGTCCAGCATCTCAGCATCTTCTTTCCACGTC 660
Db 703 GCAGCTGTCATGAAAGCCACAGGACCTTTGGTCCAGCATCTCAGCATCTTCTTTCCACGTC 762
QY 661 TGTCCAGGGTCCCTGGGAGAGACCCATCTCGGCAAAAGTGGCCCTACTATATCAACCA 720
Db 763 TGTCCAGGGTCCCTGGGAGAGACCCATCTCGGCAAAAGTGGCCCTACTATATCAACCA 822
QY 721 CGAGCTCAAAACCTTGTGGGACCATCCCAAAATGACAGAGCTTACCACTTTAGC 780
Db 823 CGAGCTCAAAACCTTGTGGGACCATCCCAAAATGACAGAGCTTACCACTTTAGC 882
QY 781 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
Db 883 TGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 942
QY 841 GAAGGCCCTTTGCTGATCTCTGAGCCTGTGAGCTGATGATGATGATGATGATGATGATGAT 900
Db 943 GAAGGCCCTTTGCTGATCTCTGAGCCTGTGAGCTGATGATGATGATGATGATGATGATGAT 1002
QY 901 CAACCTCAAGCAAAATGACAGCCCATGATATCTTCCGAGATTATTAATTTGTTGACCAC 960
Db 1003 CAACCTCAAGCAAAATGACAGCCCATGATATCTTCCGAGATTATTAATTTGTTGACCAC 1062
QY 961 TATTTATGACCGCTGAGCAAGAGCAAAATTTGGTCAAGCTCCCTCTCTGCGTGGGA 1020
Db 1063 TATTTATGACCGCTGAGCAAGAGCAAAATTTGGTCAAGCTCCCTCTCTGCGTGGGA 1122
QY 1021 TATGTCCTGAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1123 TATGTCCTGAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1182
QY 1081 CCTGTCTTTTAAACTGGCATCAATTTCCCTGTGTAAAGCAATTTGGAAGACAAAGTACAG 1140
Db 1183 CCTGTCTTTTAAACTGGCATCAATTTCCCTGTGTAAAGCAATTTGGAAGACAAAGTACAG 1242
QY 1141 ATACCTTTTCAAGCAAGTGCAGTCAACAGGATTTTGGACAGGACGAGGCTGGGCT 1200
Db 1243 ATACCTTTTCAAGCAAGTGCAGTCAACAGGATTTTGGACAGGACGAGGCTGGGCT 1302
QY 1201 CCTTCTGATGATTTCTATCCAAATTTCAAGACAGTGTGGGTGAAGTTGCAATCTTTGGGG 1260
Db 1303 CCTTCTGATGATTTCTATCCAAATTTCAAGACAGTGTGGGTGAAGTTGCAATCTTTGGGG 1362
QY 1261 CAGTAAACATTGAGCAAGTGTCCGAGCTGCTGCCAAACAAAGTGCCTTACTATATCAACCA 1301
```

Db 1363 CAGTAAACATTGAGCAAGTGTCCGAGCTGCTTCCAATTG 1403

RESULT 2

US-09-845-416-4
; Sequence 4, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-4

Query Match 100.0%; Score 1301; DB 13; Length 2169;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATTTTCCAGCAGTTTCAGAAAGCAGAACGATGATACATAGGGCCCTTCAAGAGGGAATTGAA 60
Db 451 CGATTTTCCAGCAGTTTCAGAAAGCAGAACGATGATACATAGGGCCCTTCAAGAGGGAATTGAA 510
QY 61 AACTAAAGAACCTTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGCA 120
Db 511 AACTAAAGAACCTTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGCA 570
QY 121 GCCTTTCCGAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCTCTCTGAGAGAG 180
Db 571 GCCTTTCCGAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCTCTCTGAGAGAG 630
QY 181 AGCCCAAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 240
Db 631 AGCCCAAGAAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATACTGAGTGGGA 690
QY 241 AAAATTGAACTTGCACCTCCGCTGACTGGCAGAGAAATAGATGAGACCTTCAAGAGACT 300
Db 691 AAAATTGAACTTGCACCTCCGCTGACTGGCAGAGAAATAGATGAGACCTTCAAGAGACT 750
QY 301 CCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
Db 751 CCAGGAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 810
QY 361 CAAGGGATCTCGGAGCCCGCTGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 420
Db 811 CAAGGGATCTCGGAGCCCGCTGGCGATCTCTCATTTGACTCTCTCCAAGATCACCTCGA 870
QY 421 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACTGAGCCACCTCAA 480
Db 871 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCTCTGAAAGAGAACTGAGCCACCTCAA 930
QY 481 TGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC 540
Db 931 TGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC 990
QY 541 TCTGGAAGACCTTGAACACCGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 600
Db 991 TCTGGAAGACCTTGAACACCGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 1050
QY 601 GCAGCTGATGAAGCCCAAGGACTTTTGGTCCAGCATCTCAGACATCTTCTTCCAGCTC 660
Db 1051 GCAGCTGATGAAGCCCAAGGACTTTTGGTCCAGCATCTCAGACATCTTCTTCCAGCTC 1110
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 720


```
Db 2919 ATACCTTTTCAAGCAAGTGGAGGTTTCAACAGATTTTGTGACCGAGCGCTGGGCT 2978
QY 1201 CCTTCTGCATGATTCATCAAAATCCAAAGACAGTTGGTGAAAGTTGCATCCTTTGGGG 1260
Db 2979 CCTTCTGCATGATTCATCAAAATCCAAAGACAGTTGGTGAAAGTTGCATCCTTTGGGG 3038
QY 1261 CAGTACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 1301
Db 3039 CAGTAACTTGGCCCAAGTGTCCGAGCTGCTTCCAATTG 3079

RESULT 4
US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-10

Query Match 100.0%; Score 1301; DB 13; Length 3531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 60
Db 1800 CGACTTTCCAGCAGTTTCAAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAA 1859
QY 61 AACTAAGAACCTGTAAATCATGATGACTCTTGAGACTGTACCAATATTTCTCAGACAGCA 120
Db 1860 AACTAAGAACCTGTAAATCATGATGACTCTTGAGACTGTACCAATATTTCTCAGACAGCA 1919
QY 121 GCCTTTGGAAGACTAGAGAACTTACAGAGGCCAGAGAGCTGCTCTCGAGGAGAG 180
Db 1920 GCCTTTGGAAGACTAGAGAACTTACAGAGGCCAGAGAGCTGCTCTCGAGGAGAG 1979
QY 181 AGCCAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACCTAGTGGGA 240
Db 1980 AGCCAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACCTAGTGGGA 2039
QY 241 AAAATTGAACCTGACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 300
Db 2040 AAAATTGAACCTGACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACT 2099
QY 301 CCAGGAACCTCAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCACGCTGAGTGAT 360
Db 2100 CCAGGAACCTCAGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCACGCTGAGTGAT 2159
QY 361 CAAGGATCTCTGCGAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 420
Db 2160 CAAGGATCTCTGCGAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGA 2219
QY 421 GAAAGTCAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAACGCTGAGCCAGTCAA 480
Db 2220 GAAAGTCAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAACGCTGAGCCAGTCAA 2279
QY 481 TGACCTTGTCTGCGAGCTTACCACTTTGGGCACTTCACTCTACCGTATAAATCTCAGCAC 540
Db 2280 TGACCTTGTCTGCGAGCTTACCACTTTGGGCACTTCACTCTACCGTATAAATCTCAGCAC 2339
```

```
QY 541 TCTGGAAGCCTCAACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCTAG 600
Db 2340 TCTGGAAGCCTCAACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCTAG 2399
QY 601 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCAGCATCTCAGACATCTTCTTCCAGCTC 660
Db 2400 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCAGCATCTCAGACATCTTCTTCCAGCTC 2459
QY 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 720
Db 2460 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 2519
QY 721 CGAGACTCAAAACAACTTGTGGGACCATCCAAATGACAGAGCTCTACAGTCTTTAGC 780
Db 2520 CGAGACTCAAAACAACTTGTGGGACCATCCAAATGACAGAGCTCTACAGTCTTTAGC 2579
QY 781 TGACCTGAATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGA 840
Db 2580 TGACCTGAATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACCTCCGAAGACTGA 2639
QY 841 GAAGGCCCTTTTGTCTTGGATCTCTTGGAGCCTGTGATGATGATGATGATGATGATGAT 900
Db 2640 GAAGGCCCTTTTGTCTTGGATCTCTTGGAGCCTGTGATGATGATGATGATGATGATGAT 2699
QY 901 CAACCTCAAGCAAAATGACAGCCCATGATATCTCAGAGATTTAAATTTGTTGACCAAC 960
Db 2700 CAACCTCAAGCAAAATGACAGCCCATGATATCTCAGAGATTTAAATTTGTTGACCAAC 2759
QY 961 TATTATGACCGCTCGAGCAAGACCAACAATTTGGTCAACGCTCCCTCTCTGCGTGA 1020
Db 2760 TATTATGACCGCTCGAGCAAGACCAACAATTTGGTCAACGCTCCCTCTCTGCGTGA 2819
QY 1021 TATGTGCTGAACTGGCTGCTGAAATGTTTATGATGAGGACCAAGAGGAGGATCGGT 1080
Db 2820 TATGTGCTGAACTGGCTGCTGAAATGTTTATGATGAGGACCAAGAGGAGGATCGGT 2879
QY 1081 CCTGTCTTTAAACTGGCATCAATTTCCCTGTGTAAGCACATTTTGGTAAAGACAAAGTACAG 1140
Db 2880 CCTGTCTTTAAACTGGCATCAATTTCCCTGTGTAAGCACATTTTGGTAAAGACAAAGTACAG 2939
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGCCAGCGCAGGCTGGGCT 1200
Db 2940 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGCCAGCGCAGGCTGGGCT 2999
QY 1201 CCTTCTGCATGATTCATCCAAATCCAAAGACAGTTGGTGAAAGTTGCATCCTTTGGGG 1260
Db 3000 CCTTCTGCATGATTCATCCAAATCCAAAGACAGTTGGTGAAAGTTGCATCCTTTGGGG 3059
QY 1261 CAGTAAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 1301
Db 3060 CAGTAAACATTGAGCCCAAGTGTCCGAGCTGCTTCCAATTG 3100
```

```
RESULT 5
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9
```


Query Match		100.0%;	Score 1301;	DB 13;	Length 3859;	
Best Local Similarity		100.0%;	Pred. No. 0;			
Matches 1301;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA	60			
DB	2127	CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA	2186			
QY	61	AACATAAGAACCTGTATCATGTAGTACTCTTGTAGACTGTACGAATATTTCTGACAGAGCA	120			
DB	2187	AACATAAGAACCTGTATCATGTAGTACTCTTGTAGACTGTACGAATATTTCTGACAGAGCA	2246			
QY	121	GCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAG	180			
DB	2247	GCCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGAGAG	2306			
QY	181	AGCCCAAGATGTCACTCGGCTTCTACAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA	240			
DB	2307	AGCCCAAGATGTCACTCGGCTTCTACAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA	2366			
QY	241	AAATTTGAACTGCACCTCCCTGATCGGCGATCTCTCATTTGACTCTCTCCAGAGTCACTCGA	300			
DB	2367	AAATTTGAACTGCACCTCCCTGATCGGCGATCTCTCATTTGACTCTCTCCAGAGTCACTCGA	2426			
QY	301	CCAGGAACTTCAAGAGGCCAGGATGAGCTGACCTCAAGCTGGCCCAAGCTGAGTGAT	360			
DB	2427	CCAGGAACTTCAAGAGGCCAGGATGAGCTGACCTCAAGCTGGCCCAAGCTGAGTGAT	2486			
QY	361	CAAGGATCTCTGGCAGCCCTGGGCGATCTCTCATTTGACTCTCTCCAGAGTCACTCGA	420			
DB	2487	CAAGGATCTCTGGCAGCCCTGGGCGATCTCTCATTTGACTCTCTCCAGAGTCACTCGA	2546			
QY	421	GAAAGTCAAGGCACTTCAGAGGAAATGGGCTCTGAAAGAGAACTGAGGACCACTCAA	480			
DB	2547	GAAAGTCAAGGCACTTCAGAGGAAATGGGCTCTGAAAGAGAACTGAGGACCACTCAA	2606			
QY	481	TGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC	540			
DB	2607	TGACCTTGCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCAC	2666			
QY	541	TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTCGAGAACGAGTCAAG	600			
DB	2667	TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTCGAGAACGAGTCAAG	2726			
QY	601	GCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGGATCTCAGCACTTTCTTTCCACGTC	660			
DB	2727	GCAGCTGCATGAAGCCACAGGACCTTTGGTCCAGGATCTCAGCACTTTCTTTCCACGTC	2786			
QY	661	TGTCAGGAGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA	720			
DB	2787	TGTCAGGAGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA	2846			
QY	721	CGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC	780			
DB	2847	CGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC	2906			
QY	781	TGACCTGAATTAATGTGAGATTTCTGAGCTTATAGACTGCCATGAACTCCGAAAGACTGCA	840			
DB	2907	TGACCTGAATTAATGTGAGATTTCTGAGCTTATAGACTGCCATGAACTCCGAAAGACTGCA	2966			
QY	841	GAAGGCCCTTTGCTTGGATCTTTGAGCCCTGTGAGTGCATGTGATGCCCTTGGACAGCA	900			
DB	2967	GAAGGCCCTTTGCTTGGATCTTTGAGCCCTGTGAGTGCATGTGATGCCCTTGGACAGCA	3026			
QY	901	CAACCTCAAGCAAAATGACAGCCCAATGGATATCCCTGCAGATTTAATTTGTTTGACAC	960			
DB	3027	CAACCTCAAGCAAAATGACAGCCCAATGGATATCCCTGCAGATTTAATTTGTTTGACAC	3086			
QY	961	TATTTATGACCCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGA	1020			
DB	3087	TATTTATGACCCCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGTGA	3146			
Query Match		100.0%;	Score 1301;	DB 13;	Length 3999;	
Best Local Similarity		100.0%;	Pred. No. 0;			
Matches 1301;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA	60			
DB	2268	CGACTTTCCAGCAGTTCCAGAGCAGACGATGTACATAGGGCCCTTCAAGAGGGAATTGAA	2327			
QY	61	AACATAAGAACCTGTATCATGTAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA	120			
DB	2328	AACATAAGAACCTGTATCATGTAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCA	2387			
QY	121	GCCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGGAGAG	180			
DB	2388	GCCTTTGGAAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGGAGAG	2447			
QY	181	AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA	240			
DB	2448	AGCCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA	2507			
QY	241	AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT	300			
DB	2508	AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT	2567			
QY	301	CCAGGAACTTCAAGAGGCCAGGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGAT	360			
DB	2568	CCAGGAACTTCAAGAGGCCAGGATGAGCTGGAACCTCAAGCTGCGCAAGCTGAGGTGAT	2627			
QY	361	CAAGGGAATCTTGGCAGAGCCCGTGGGCGGATCTCTCATTTGACTCTCTCCTCAAGATCACCTCGA	420			
DB	2628	CAAGGGAATCTTGGCAGAGCCCGTGGGCGGATCTCTCATTTGACTCTCTCCTCAAGATCACCTCGA	2687			

RESULT 6

US-09-845-416-6

; Sequence 6, Application US/09845416

; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO

; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

; FILE REFERENCE: DEL142

; CURRENT APPLICATION NUMBER: US/09/845,416

; PRIOR FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 6

; LENGTH: 3999

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-845-416-6

```

; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match      100.0%; Score 1301; DB 13; Length 4182;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTTCAGAACGACGATGATGAGGCTTCAAGAGGGAATTTGAA 60
DB 2451 CGACTTTCCAGCAGTTTCAGAACGACGATGATGAGGCTTCAAGAGGGAATTTGAA 2510

QY 61 AACTTAAGAACCTTGAATCATGATGATCTTGTAGAGCTGTACGAAATATTTCTGACAGAGCA 120
DB 2511 AACTTAAGAACCTTGAATCATGATGATCTTGTAGAGCTGTACGAAATATTTCTGACAGAGCA 2570

QY 121 GCCTTTTGGAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCTTCTGAGGAG 180
DB 2571 GCCTTTTGGAGGACTAGAGAACTCTACAGGAGGCCAGAGAGCTGCTTCTGAGGAG 2630

QY 181 AGCCCAAGATGTCACTCGGCTTCTAGCAAGCAGGCTGAGGAGGTCAATCTAGTGGGA 240
DB 2631 AGCCCAAGATGTCACTCGGCTTCTAGCAAGCAGGCTGAGGAGGTCAATCTAGTGGGA 2690

QY 241 AAAATTGAACTCGACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTCAAAAGACT 300
DB 2691 AAAATTGAACTCGACTCCGCTGACTGGCAGAGAAATAGATGAGACCTTCAAAAGACT 2750

QY 301 CCAGGAATTTCAAGAGCCACGATGATGCTGAGCTCAAGCTGGCCCAAGCTGAGTGAT 360
DB 2751 CCAGGAATTTCAAGAGCCACGATGATGCTGAGCTCAAGCTGGCCCAAGCTGAGTGAT 2810

QY 361 CAAGGATCTGGCAGGCGCTGGCGGATCTCTCATTTGACTCTCTCCAGATCACCTCGA 420
DB 2811 CAAGGATCTGGCAGGCGCTGGCGGATCTCTCATTTGACTCTCTCCAGATCACCTCGA 2870

QY 421 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTGAAAGAGAACTGAGGACCTCAA 480
DB 2871 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCTCTGAAAGAGAACTGAGGACCTCAA 2930

QY 481 TGACCTTGTCTGCGCAGCTTACCACTTTTGGGCACTTCAAGCTCTCAAGCTTAACTCAGCAC 540
DB 2931 TGACCTTGTCTGCGCAGCTTACCACTTTTGGGCACTTCAAGCTCTCAAGCTTAACTCAGCAC 2990

QY 541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCA 600
DB 2991 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCA 3050

QY 601 GCAGCTGCATGAAGCCACAGGAGCTTTGGTTCAGGATCTCAGCACTTTCTTTCCAGCTC 660
DB 3051 GCAGCTGCATGAAGCCACAGGAGCTTTGGTTCAGGATCTCAGCACTTTCTTTCCAGCTC 3110

QY 661 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCTTCTATATCAACCA 720
DB 3111 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCTTCTATATCAACCA 3170

QY 721 CGAGACTCAAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAC 780
DB 3171 CGAGACTCAAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAC 3230

QY 781 TGACCTGAATATGTCTCAGATTTCTCAGCTTATAGACTGCCATGAAATCCGAGAGCTGCA 840
DB 3231 TGACCTGAATATGTCTCAGATTTCTCAGCTTATAGACTGCCATGAAATCCGAGAGCTGCA 3290

QY 841 GAAGGCCCTTTGCTGGATCTTTGAGGCTGTGAGTGCATGTGATGCTTGGACAGCA 900
DB 3291 GAAGGCCCTTTGCTGGATCTTTGAGGCTGTGAGTGCATGTGATGCTTGGACAGCA 3350

421 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCTCTGAAAGAGAACTGAGGACCTCAA 480
2688 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCTCTGAAAGAGAACTGAGGACCTCAA 2747
481 TGACCTTGTCTGCGCAGCTTACCACTTTTGGGCACTTCAAGCTTCAAGCTCAGCAC 540
2748 TGACCTTGTCTGCGCAGCTTACCACTTTTGGGCACTTCAAGCTTCAAGCTCAGCAC 2807
541 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCA 600
2808 TCTGGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCA 2867
601 GCAGCTGCATGAAGCCACAGGAGCTTTGGTTCAGGATCTCAGCACTTTCTTTCCAGCTC 660
2868 GCAGCTGCATGAAGCCACAGGAGCTTTGGTTCAGGATCTCAGCACTTTCTTTCCAGCTC 2927
661 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCTTCTATATCAACCA 720
2928 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGGCTTCTATATCAACCA 2987
721 CGAGACTCAAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAC 780
2988 CGAGACTCAAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACCACTTTTAC 3047
781 TGACCTGAATATGTCTCAGATTTCTCAGCTTATAGGACTGGCATGAACTCCGAGAGCTGCA 840
3048 TGACCTGAATATGTCTCAGATTTCTCAGCTTATAGGACTGGCATGAACTCCGAGAGCTGCA 3107
841 GAAGGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGCTGATGCTTGGAGCCAGCA 900
3108 GAAGGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGCTGATGCTTGGAGCCAGCA 3167
901 CAACCTCAAGCAAAATGACAGGCGCATGGAATCTTGGAGATTTAATTTGTTGACCA 960
3168 CAACCTCAAGCAAAATGACAGGCGCATGGAATCTTGGAGATTTAATTTGTTGACCA 3227
961 TATTTATGACGCTGGAGCAAGACAGCAACAATTTGGTCAAGCTCCCTCTCTGGGTGA 1020
3228 TATTTATGACGCTGGAGCAAGACAGCAACAATTTGGTCAAGCTCCCTCTCTGGGTGA 3287
1021 TATGCTGTGAATCTGCTGATGTTATGATACGGAGCAAGACAGGAGGATCGGT 1080
3288 TATGCTGTGAATCTGCTGATGTTATGATACGGAGCAAGACAGGAGGATCGGT 3347
1081 CCTGTCTTTAAACTGGGATCATTTCCCTGTGTAAGCAGCATTTTGGAGAGCAAGTACAG 1140
3348 CCTGTCTTTAAACTGGGATCATTTCCCTGTGTAAGCAGCATTTTGGAGAGCAAGTACAG 3407
1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGAGGCTGGGCT 1200
3408 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGAGGCTGGGCT 3467
1201 CCTTCTGCATGATTTCTATCCAAATTCAGACAGATTTGGGTGAAGTTGCACTTTGGGG 1260
3468 CCTTCTGCATGATTTCTATCCAAATTCAGACAGATTTGGGTGAAGTTGCACTTTGGGG 3527
1261 CAGTAACTTGAAGCAAGTGTCCGAGCTGCTTCCAAATTTG 1301
3528 CAGTAACTTGAAGCAAGTGTCCGAGCTGCTTCCAAATTTG 3568

RESULT 7
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE OF INVENTION: DEL142
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
```

QY 901 CAACCTCAAGCAAAATGACCAAGCCATGGATATCCTGCAGATTATTAATTGTTTGACCC 960
DB 3351 CAACCTCAAGCAAAATGACCAAGCCATGGATATCCTGCAGATTATTAATTGTTTGACCC 3410
QY 961 TATTATGACCGCTCGAGCAGAGCAACAATTTGGTCAAGTCCCTCTCTCGCGTGA 1020
DB 3411 TATTATGACCGCTCGAGCAGAGCAACAATTTGGTCAAGTCCCTCTCTCGCGTGA 3470
QY 1021 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGGATCCGTGT 1080
DB 3471 TATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGAGCAACAGGAGGATCCGTGT 3530
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTGAAGCAATTTGGAAGCAAAATGACG 1140
DB 3531 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTGAAGCAATTTGGAAGCAAAATGACG 3590
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTCAACAGGATTTTGTACCGAGCGAGGCTGGGCT 1200
DB 3591 ATACCTTTTCAAGCAAGTGGCAAGTCAACAGGATTTTGTACCGAGCGAGGCTGGGCT 3650
QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCAAGACAGTGGGTGAAGTTGCATCTTTGGGG 1260
DB 3651 CCTTCTGCATGATTTCTATCCAAATTTCAAGACAGTGGGTGAAGTTGCATCTTTGGGG 3710
QY 1261 CAGTAACATTGAGCCAGTGTCCGGAGCTGCTTCCAAATTG 1301
DB 3711 CAGTAACATTGAGCCAGTGTCCGGAGCTGCTTCCAAATTG 3751

RESULT 8
US-09-845-416-31
; Sequence 31, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 4476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-31

Query Match 100.0%; Score 1301; DB 13; Length 4476;
Best Local Similarity 100.0%; Pred. No 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTCCAGAGCAAGACGATGATACATAGGCTTCCAGAGGGAATTGAA 60
DB 2535 CGACTTTCCAGCAGTTCCAGAGCAAGACGATGATACATAGGCTTCCAGAGGGAATTGAA 2594
QY 61 AACTTAAGACCTGTAATCATGATGATCTTTGAGACTGTACGATATTTCTGACAGCA 120
DB 2595 AACTTAAGACCTGTAATCATGATGATCTTTGAGACTGTACGATATTTCTGACAGCA 2654
QY 121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCCTCAGGAGAG 180
DB 2655 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCCTCAGGAGAG 2714
QY 181 AGCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATCTAGTGGGA 240
DB 2715 AGCCAGAAATGTACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATCTAGTGGGA 2774
QY 241 AARATTGACCTGACCTCGCTGCTGCGAGAGAAAATAGATGAGACCTCTCAAGAGCT 300

RESULT 9

US-09-845-416-30

DB 2775 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTTGAAAGACT 2834
QY 301 CCAGGAACCTTCAAGAGGCCACCGATGAGTGCACCTCAAGCTGCGCAAGCTGAGGTGAT 360
DB 2835 CCAGGAACCTTCAAGAGGCCACCGATGAGTGCACCTCAAGCTGCGCAAGCTGAGGTGAT 2894
QY 361 CAAGGGATTCCTGGCAGCCGCTGGCGGATCTCTCATTTGACTCTCTCCAGATCACTTCGA 420
DB 2895 CAAGGGATTCCTGGCAGCCGCTGGCGGATCTCTCATTTGACTCTCTCCAGATCACTTCGA 2954
QY 421 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGCGCTCTGAAAGAGAACTGAGCCACGTCAC 480
DB 2955 GAAAGTCAAGGCACCTTCGAGGAGAAATTTGCGCTCTGAAAGAGAACTGAGCCACGTCAC 3014
QY 481 TGAACCTTCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACGATATAACCTCAGCAC 540
DB 3015 TGAACCTTCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACGATATAACCTCAGCAC 3074
QY 541 TCTGGAACACCTGGAACACCAAGCTTCTGAGTGGCGCTTCGAGGACCGAGTCAG 600
DB 3075 TCTGGAACACCTGGAACACCAAGCTTCTGAGTGGCGCTTCGAGGACCGAGTCAG 3134
QY 601 GCAGTGCATGAAGCCCAACAGGACTTTTGTCCAGCATCTCAGCACTTTCTTTCCAGTTC 660
DB 3135 GCAGTGCATGAAGCCCAACAGGACTTTTGTCCAGCATCTCAGCACTTTCTTTCCAGTTC 3194
QY 661 TGTCCAGGCTCCTGGGAGAGAGCCATCTGCCAAAACAAAGTGGCCCTACTATATCAACCA 720
DB 3195 TGTCCAGGCTCCTGGGAGAGAGCCATCTGCCAAAACAAAGTGGCCCTACTATATCAACCA 3254
QY 721 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 780
DB 3255 CGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGC 3314
QY 781 TGACCTGAATTAATGTAGATTTCTCAGCTTATAGACTGCGCATGAACTCCGAAGACTGCA 840
DB 3315 TGACCTGAATTAATGTAGATTTCTCAGCTTATAGACTGCGCATGAACTCCGAAGACTGCA 3374
QY 841 GAAGCCCTTTGCTTGTGATCTCTTGTAGCTGTCTGAGTGCATGTGATGCTTGGACCAAGCA 900
DB 3375 GAAGCCCTTTGCTTGTGATCTCTTGTAGCTGTCTGAGTGCATGTGATGCTTGGACCAAGCA 3434
QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTAATTAATTGTTGACCA 960
DB 3435 CAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAGATTAATTAATTGTTGACCA 3494
QY 961 TATTTATGACCGCTGGGAGCAAGCAACAATTTGCTCAACCTCCCTCTCTCGGTGGA 1020
DB 3495 TATTTATGACCGCTGGGAGCAAGCAACAATTTGCTCAACCTCCCTCTCTCGGTGGA 3554
QY 1021 TATGTGCTGAACCTGGCTGCTGATGTTTATGATACGGGAGCAACAGGAGGATCCGTGT 1080
DB 3555 TATGTGCTGAACCTGGCTGCTGATGTTTATGATACGGGAGCAACAGGAGGATCCGTGT 3614
QY 1081 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAACAGTACAG 1140
DB 3615 CCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAACAGTACAG 3674
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCGAGGCTGGGCT 1200
DB 3675 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCGAGGCTGGGCT 3734
QY 1201 CCTTCTGCATGATTTCTATCCAAATTTCAAGACAGTGTGGGTGAAGTTCATCTTTGGGG 1260
DB 3735 CCTTCTGCATGATTTCTATCCAAATTTCAAGACAGTGTGGGTGAAGTTCATCTTTGGGG 3794
QY 1261 CAGTAACATTGAGCCAGTGTCCGGAGCTGCTTCCAAATTG 1301
DB 3795 CAGTAACATTGAGCCAGTGTCCGGAGCTGCTTCCAAATTG 3835

```
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4498
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-845-416-30

Query Match      100.0%; Score 1301; DB 13; Length 4498;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTCCAGCGATTCCAGAGCAGAACGATGTACATAGGGCCCTTCAGAGGGAATTGAA 60
DB 2557 CGACTTCCAGCGATTCCAGAGCAGAACGATGTACATAGGGCCCTTCAGAGGGAATTGAA 2616
QY 61 AACTAAGAACCTGTAACTCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
DB 2617 AACTAAGAACCTGTAACTCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2676
QY 121 GCCTTTGGAAGGACTAGAGAACTTACACAGAGGCCACAGAGAGCTCCCTCGAGGAGAG 180
DB 2677 GCCTTTGGAAGGACTAGAGAACTTACACAGAGGCCACAGAGAGCTCCCTCGAGGAGAG 2736
QY 181 AGCCCAAGTGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 240
DB 2737 AGCCCAAGTGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATACCTGAGTGGGA 2796
QY 241 AAATTAAGCTGACCTCCCTGACTGGCAGAGAGAAATAGATGAGACCTTGAAGACT 300
DB 2797 AAATTAAGCTGACCTCCCTGACTGGCAGAGAGAAATAGATGAGACCTTGAAGACT 2856
QY 301 CCAGGAACTTCAAGAGCCACGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGTGTAT 360
DB 2857 CCAGGAACTTCAAGAGCCACGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGTGTAT 2916
QY 361 CAAGGATCTCGGAGCCCGTGGCGATCTCTCAATTGACTCTCTCCAAAGATCACTCGA 420
DB 2917 CAAGGATCTCGGAGCCCGTGGCGATCTCTCAATTGACTCTCTCCAAAGATCACTCGA 2976
QY 421 GAAAGTCAAGCACTTCGAGAGAAATGCGCTCTGAAAGAGAAAGTGAAGCAAGTCA 480
DB 2977 GAAAGTCAAGCACTTCGAGAGAAATGCGCTCTGAAAGAGAAAGTGAAGCAAGTCA 3036
QY 481 TGACTTGTCTCGCCAGCTTACCCTTTGGCCATTCCAGCTCTCACCGTATAACCTCAGCAC 540
DB 3037 TGACTTGTCTCGCCAGCTTACCCTTTGGCCATTCCAGCTCTCACCGTATAACCTCAGCAC 3096
QY 541 TCTGGAAGACTGAACACACAGATGGAAGCTTCTGAGGTGGCGTGGAGGACCGAGTCA 600
DB 3097 TCTGGAAGACTGAACACACAGATGGAAGCTTCTGAGGTGGCGTGGAGGACCGAGTCA 3156
QY 601 CGAGTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCATCTTCTTCCACGTC 660
DB 3157 CGAGTGCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCATCTTCTTCCACGTC 3216
QY 661 TGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTCTATATCAACCA 720
DB 3217 TGTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTCTATATCAACCA 3276
QY 721 CGAGACTCAACAACTGCTGGGACCATCCCAATTCACAGAGCTCTACCAAGCTTTAGC 780
DB 3276 CGAGACTCAACAACTGCTGGGACCATCCCAATTCACAGAGCTCTACCAAGCTTTAGC 3336

3277 CGAGACTCAACAACTGCTGGGACCATCCCAATTCACAGAGCTCTACCAAGCTTTAGC 3336
QY 781 TGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 840
DB 3337 TGACCTGAATAATGTGAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCA 3396
QY 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCTGTGTCAGCTGATGATGCTTGGACAGCA 900
DB 3397 GAAGGCCCTTTGCTTGGATCTCTTGAGCTGTGTCAGCTGATGATGCTTGGACAGCA 3456
QY 901 CAACCTCAAGCAAAATGACCCAGCCATGGATATCTTCGAGATTTATTAAATTGTTTGACCA 960
DB 3457 CAACCTCAAGCAAAATGACCCAGCCATGGATATCTTCGAGATTTATTAAATTGTTTGACCA 3516
QY 961 TATTATGACCGCTCGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTCGCTGGA 1020
DB 3517 TATTATGACCGCTCGAGCAAGAGCACAACAATTTGGTCAACGTCCTCTCTCGCTGGA 3576
QY 1021 TATGTGCTGAAGTGGCTGCTGAATGTTTATGATACGGGACGAGCAAGGAGGATCCGTGT 1080
DB 3577 TATGTGCTGAAGTGGCTGCTGAATGTTTATGATACGGGACGAGCAAGGAGGATCCGTGT 3636
QY 1081 CCTGCTCTTTAAATACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 1140
DB 3637 CCTGCTCTTTAAATACTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAG 3696
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACAGCGCAGGCTGGCCCT 1200
DB 3697 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACAGCGCAGGCTGGCCCT 3756
QY 1201 CCTTCTGCATGATTTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCTTTGGGGG 1260
DB 3757 CCTTCTGCATGATTTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCTTTGGGGG 3816
QY 1261 CAGTAAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 1301
DB 3817 CAGTAAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG 3857

RESULT 10
US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 29
; LENGTH: 4825
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-845-416-29

Query Match      100.0%; Score 1301; DB 13; Length 4825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCACAGCTTCAGAGCAGAAACGATGTACATAGGGCCCTTCAGAGGGAATTGAA 60
DB 2884 CGACTTTCACAGCTTCAGAGCAGAAACGATGTACATAGGGCCCTTCAGAGGGAATTGAA 2943
QY 61 AACTAAGAACTGTAATCATGAGTACTCTTGAGACTGTAGCAATATTTCTGACAGAGCA 120
DB 2944 AACTAAGAACTGTAATCATGAGTACTCTTGAGACTGTAGCAATATTTCTGACAGAGCA 3003
QY 121 GCCTTTGGAAGGACTAGAGAACTCTTACAGGAGGCCACAGAGCTCTCTCGAGGAGAG 180
```

Db 3004 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAG 3063
QY 181 AGCCAGAGATGCTACTGGCTTCTACGAAGCAGGCTGAGAGGTCAATCTAGTGGGA 240
Db 3064 AGCCAGAGATGCTACTGGCTTCTACGAAGCAGGCTGAGAGGTCAATCTAGTGGGA 3123
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGCGCAGAGAAAAATAGATCAGACCTTCAAAGACT 300
Db 3124 AAAATTGAACCTGCACCTCCGCTGACTGCGCAGAGAAAAATAGATCAGACCTTCAAAGACT 3183
QY 301 CCAGGAACCTCAAGAGCCACCGATGAGCTGAGCCTCAAGCTGCGCCCAAGCTGAGTGTAT 360
Db 3184 CCAGGAACCTCAAGAGCCACCGATGAGCTGAGCCTCAAGCTGCGCCCAAGCTGAGTGTAT 3243
QY 361 CAAGGGATCTCGGAGCCGCGGCGGATCTCTCATTTGACTCTCTCCAAAGTCAACCTCGA 420
Db 3244 CAAGGGATCTCGGAGCCGCGGCGGATCTCTCATTTGACTCTCTCCAAAGTCAACCTCGA 3303
QY 421 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 480
Db 3304 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 3363
QY 481 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCCCGTATTAACCTCAGCAC 540
Db 3364 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCCCGTATTAACCTCAGCAC 3423
QY 541 TCTGGAAGACCTGAAACCCAGATGGAAGCTTCTGAGGTGGCGCTCGAGACCGAGTTCAG 600
Db 3424 TCTGGAAGACCTGAAACCCAGATGGAAGCTTCTGAGGTGGCGCTCGAGACCGAGTTCAG 3483
QY 601 GCAGCTGATGAAGCCACAGGACTTTGGTCCAGCATCTCAGACCTTCTTTCCAGCTC 660
Db 3484 GCAGCTGATGAAGCCACAGGACTTTGGTCCAGCATCTCAGACCTTCTTTCCAGCTC 3543
QY 661 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCAGAAACAAAGTGCCTTACTATATCAACCA 720
Db 3544 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCAGAAACAAAGTGCCTTACTATATCAACCA 3603
QY 721 CGAGACTCAAACTGCTGGGACCAATCCGAAATGACAGAGCTCTACAGTCTTTAGC 780
Db 3604 CGAGACTCAAACTGCTGGGACCAATCCGAAATGACAGAGCTCTACAGTCTTTAGC 3663
QY 781 TGACCTGATTAATGCTAGATCTCAGCTTATAGAGCTGCCATGAATCTCGAAGAGCTGCA 840
Db 3664 TGACCTGATTAATGCTAGATCTCAGCTTATAGAGCTGCCATGAATCTCGAAGAGCTGCA 3723
QY 841 GAAGGCCCTTTGCTTGGATCTTTGACCTGTCTAGCTGCAATGTGATGCTTGGACCAAGCA 900
Db 3724 GAAGGCCCTTTGCTTGGATCTTTGACCTGTCTAGCTGCAATGTGATGCTTGGACCAAGCA 3783
QY 901 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGCGATTAATTAATTTGTTGACCA 960
Db 3784 CAACCTCAAGCAAAATGACAGCCCATGGATATCTCGCGATTAATTAATTTGTTGACCA 3843
QY 961 TAATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAGCTCCCTCTCTCGGTGGA 1020
Db 3844 TAATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAGCTCCCTCTCTCGGTGGA 3903
QY 1021 TATGTGCTGAACCTGGCTGCTGATGTTTATGATACGGGACGACAGGAGGATCGTGT 1080
Db 3904 TATGTGCTGAACCTGGCTGCTGATGTTTATGATACGGGACGACAGGAGGATCGTGT 3963
QY 1081 CTTGTCTTTTAAACTGGCATCTTTCCCTGTGTAAAGCACTTTGGAAGCAAGTACAG 1140
Db 3964 CTTGTCTTTTAAACTGGCATCTTTCCCTGTGTAAAGCACTTTGGAAGCAAGTACAG 4023
QY 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGGCTGGGCT 1200
Db 4024 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACAGCGCAGGCTGGGCT 4083
QY 1201 CTTCTGATGATCTATCCAAATTCAGAGACAGTGGTGAAGTTCATCTTTGGGG 1260

Db 4084 CCTTCTGCATGATTCTATCAAAATTCAGAGACAGTGGGTGAAGTTCATCTTTGGGG 4143
QY 1261 CAGTAACTTGGCCCAAGTGTCCGAGCTGCTTCCAATTTG 1301
Db 4144 CAGTAACTTGGCCCAAGTGTCCGAGCTGCTTCCAATTTG 4184
RESULT 11
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35
Query Match 100.0%; Score 1301; DB 13; Length 4848;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACTTTCAGAGCTTCAGAGAGCAGATGTACATAGGGCCCTTCAAGAGGGAAATTGAA 60
Db 2907 CGACTTTCAGAGCTTCAGAGAGCAGATGTACATAGGGCCCTTCAAGAGGGAAATTGAA 2966
QY 61 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGCACAGACA 120
Db 2967 AACTAAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTCTGCACAGACA 3026
QY 121 GCCTTTGGAAGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAG 180
Db 3027 GCCTTTGGAAGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAG 3086
QY 181 AGCCAGAGATGTCACTCGGCTTCTACGAAACAGCTGAGAGGTCAATCTGAGTGGGA 240
Db 3087 AGCCAGAGATGTCACTCGGCTTCTACGAAACAGCTGAGAGGTCAATCTGAGTGGGA 3146
QY 241 AAAATTGAACCTGCACCTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCTTGAAGACT 300
Db 3147 AAAATTGAACCTGCACCTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCTTGAAGACT 3206
QY 301 CCAGGAACCTCAAGAGCCACCGATGAGCTGGAATCAAGCTGCGCAAGCTGAGGTGAT 360
Db 3207 CCAGGAACCTCAAGAGCCACCGATGAGCTGGAATCAAGCTGCGCAAGCTGAGGTGAT 3266
QY 361 CAAGGGATCTCGGAGCCGCGTGGGCGATCTCTCATTTGACTCTCTCCAAAGTCAACCTCGA 420
Db 3267 CAAGGGATCTCGGAGCCGCGTGGGCGATCTCTCATTTGACTCTCTCCAAAGTCAACCTCGA 3326
QY 421 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 480
Db 3327 GAAAGTCAAGGCACTTCGAGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCAA 3386
QY 481 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCCCGTATTAACCTCAGCAC 540
Db 3387 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCCCGTATTAACCTCAGCAC 3446
QY 541 TCTGGAAGACCTGAAACCCAGATGGAAGTTCGAGGTGGCGCTCGAGGACCGAGTTCAG 600
Db 3447 TCTGGAAGACCTGAAACCCAGATGGAAGTTCGAGGTGGCGCTCGAGGACCGAGTTCAG 3506
QY 601 GCAGCTGATGATGAGCCACAGGACTTTGGTCCAGCATCTCAGCACCTTCTTTCCAGCTC 660

Db 3507 CGAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCACTCAGCACTTCTTTCACGTC 3566
Qy 661 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGGCAAAACAAAGTGCCCTACTATATCAACA 720
Db 3567 TGTCCAGGGTCCCTGGGAGAGAGCCATCTCGGCAAAACAAAGTGCCCTACTATATCAACA 3626
Qy 721 CGAGACTCAAAACAACTTGTCTGGAGCCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 780
Db 3627 CGAGACTCAAAACAACTTGTCTGGAGCCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGC 3686
Qy 781 TGACCTGAATAATGTAGATTCAGATTCAGCTTATAGGATGCGCATGAACTCCGGAAGACTGCA 840
Db 3687 TGACCTGAATAATGTAGATTCAGATTCAGCTTATAGGATGCGCATGAACTCCGGAAGACTGCA 3746
Qy 841 GAAGGCCCTTGTCTGGATCTCTTGAGCTGTGAGCTGATGATGCTTGCAGCAGCA 900
Db 3747 GAAGGCCCTTGTCTGGATCTCTTGAGCTGTGAGCTGATGATGCTTGCAGCAGCA 3806
Qy 901 CAACCTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTAATTAATTTGTTGACCA 960
Db 3807 CAACCTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTAATTAATTTGTTGACCA 3866
Qy 961 TATTTATGACCGCTGGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGGTGA 1020
Db 3867 TATTTATGACCGCTGGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGGTGA 3926
Qy 1021 TATGTGTCTGAATCTGCTGATCTGTAATCTTATGATACGGAGCAACAGGAGGATCCGTGT 1080
Db 3927 TATGTGTCTGAATCTGCTGATCTGTAATCTTATGATACGGAGCAACAGGAGGATCCGTGT 3986
Qy 1081 CTTGTCTTTAAACTGGGATCTATCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAG 1140
Db 3987 CTTGTCTTTAAACTGGGATCTATCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAG 4046
Qy 1141 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACAGCGCAGCTGGGCT 1200
Db 4047 ATACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACAGCGCAGCTGGGCT 4106
Qy 1201 CTTTCTGATGATCTATCAAAATTCAGAGCTTGAAGTGAAGTGCATCTTTGGGGG 1260
Db 4107 CTTTCTGATGATCTATCAAAATTCAGAGCTTGAAGTGAAGTGCATCTTTGGGGG 4166
Qy 1261 CAGTAACTGAGCAGCTGCTCGGAGCTCTTCCAAATTTG 1301
Db 4167 CAGTAACTGAGCAGCTGCTCGGAGCTCTTCCAAATTTG 4207

RESULT 12

US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-28

Query Match 100.0%; Score 1301; DB 13; Length 4966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGCTTCAAGAGGAAATTGAA 60
Db 3025 CGACTTTCAGCAGTTTCAAGAGCAGAACGATGTACATAGGCTTCAAGAGGAAATTGAA 3084
Qy 61 AACTAAAGAACCTGTATATGATGATCTCTTGAGAGCTGTACGATATTTCTGACAGCA 120
Db 3085 AACTAAAGAACCTGTATATGATGATCTCTTGAGAGCTGTACGATATTTCTGACAGCA 3144
Qy 121 GCCTTTTGAAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTTGAGGAG 180
Db 3145 GCCTTTTGAAGGACTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCTCTTGAGGAG 3204
Qy 181 AGCCAGAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCTGAGTGGGA 240
Db 3205 AGCCAGAGATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCTGAGTGGGA 3264
Qy 241 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 300
Db 3265 AAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACT 3324
Qy 301 CGAGGAATTTCAAGAGGCCACCGATGAGCTGGAACCTCAAGCTGGCCAGAGCTGAGTGT 360
Db 3325 CGAGGAATTTCAAGAGGCCACCGATGAGCTGGAACCTCAAGCTGGCCAGAGCTGAGTGT 3384
Qy 361 CRAAGGATCTGCGAGCCCTGGGCGATCTCCTCATTGACTCTCTCAAGATCACCTCGA 420
Db 3385 CRAAGGATCTGCGAGCCCTGGGCGATCTCCTCATTGACTCTCTCAAGATCACCTCGA 3444
Qy 421 GAAAGTCAAGGCACTTCGAGGAGAAAATTCGCGCTCTGAAAGAGAACTGAGCACTCAA 480
Db 3445 GAAAGTCAAGGCACTTCGAGGAGAAAATTCGCGCTCTGAAAGAGAACTGAGCACTCAA 3504
Qy 481 TGACCTTGTCTGCGAGCTTACCACTTTGGGCATTCAGCTCTCAACCTATTAACCTCAGCA 540
Db 3505 TGACCTTGTCTGCGAGCTTACCACTTTGGGCATTCAGCTCTCAACCTATTAACCTCAGCA 3564
Qy 541 TCTGGAAGACTCTGAAACACAGATGGAAGCTTCTGCGAGGTGGCCGTGAGGACCGAGT 600
Db 3565 TCTGGAAGACTCTGAAACACAGATGGAAGCTTCTGCGAGGTGGCCGTGAGGACCGAGT 3624
Qy 601 GCAGCTGCATGAAGCCACAGGAGCTTTGGTTCAGATCTCAGACATCTTCTTCCAGCTC 660
Db 3625 GCAGCTGCATGAAGCCACAGGAGCTTTGGTTCAGATCTCAGACATCTTCTTCCAGCTC 3684
Qy 661 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTGCCTTACTATATCAACCA 720
Db 3685 TGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTGCCTTACTATATCAACCA 3744
Qy 721 CGAGACTCAAAACAACTTGTCTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 780
Db 3745 CGAGACTCAAAACAACTTGTCTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGC 3804
Qy 781 TGACCTGAATAATGTCAAGATTCAGCTTATAGGATGCGCATGAACTCCGGAAGACTGCA 840
Db 3805 TGACCTGAATAATGTCAAGATTCAGCTTATAGGATGCGCATGAACTCCGGAAGACTGCA 3864
Qy 841 GAAGGCCCTTTGTCTGGATCTCTTGAGCCCTGTGAGTGCATGTGATGCTTGGACAGCA 900
Db 3865 GAAGGCCCTTTGTCTGGATCTCTTGAGCCCTGTGAGTGCATGTGATGCTTGGACAGCA 3924
Qy 901 CAACCTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTAATTAATTTGTTGACCA 960
Db 3925 CAACCTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTAATTAATTTGTTGACCA 3984
Qy 961 TATTTATGACCGCTGGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 1020
Db 3985 TATTTATGACCGCTGGAGCAAGAGCACAACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 4044
Qy 1021 TATGTGTCTGAACCTGGCTGTCTGATATGATACGGGAGCAACAGGAGGATCCGTGT 1080
Db 4045 TATGTGTCTGAACCTGGCTGTCTGATATGATACGGGAGCAACAGGAGGATCCGTGT 4104
Qy 1081 CTTGTCTTTTAAACTGGCAGCTATTTCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAG 1140

Db 4105 CCTGCTTTTAAAGCTGGCATCTTCCCTGTGTAAAGACACATTTTGAAGACAAAGTACAG 4164
Qy 1141 ATACCTTTTCAAGCAGTGGCAGTTCACAGAGATTTTGTGACAGCGGAGGCTGGGCT 1200
Db 4165 ATACCTTTTCAAGCAGTGGCAGTTCACAGAGATTTTGTGACAGCGGAGGCTGGGCT 4224
Qy 1201 CCTTCTGCATGATTTCTATCCAAATTCACAGACAGTGGGTGAAGTTGCCATCCCTTTGGGG 1260
Db 4225 CCTTCTGCATGATTTCTATCCAAATTCACAGACAGTGGGTGAAGTTGCCATCCCTTTGGGG 4284
Qy 1261 CAGTAACTTGAAGCAAGTTCGAGAGCTGCTTCCAAATTG 1301
Db 4285 CAGTAACTTGAAGCAAGTTCGAGAGCTGCTTCCAAATTG 4325

RESULT 13
US-09-845-416-34
; Sequence 34, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR FILING DATE: 60/200,777
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 4990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-34

Query Match 100.0%; Score 1301; DB 13; Length 4990;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACTTTCCAGCAGTTCAGAGCAGACGATGATACATAGGCGCTTCAAGAGGGAATTGAA 60
Db 3049 CGACTTTCCAGCAGTTCAGAGCAGACGATGATACATAGGCGCTTCAAGAGGGAATTGAA 3108
Qy 61 AACTAAAGAACCTGTATCATGTAGTACTCTTGAGACTGACGATATTTCTGACAGCA 120
Db 3109 AACTAAAGAACCTGTATCATGTAGTACTCTTGAGACTGACGATATTTCTGACAGCA 3168
Qy 121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAG 180
Db 3169 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCCTGAGGAG 3228
Qy 181 AGCCAGAAATGTACTCGGTTTACGAAAGAGGCTGAGAGGTCAATCTAGTGGGA 240
Db 3229 AGCCAGAAATGTACTCGGTTTACGAAAGAGGCTGAGAGGTCAATCTAGTGGGA 3288
Qy 241 AATATTGAACTGCACTCCCTGACTGCGAGAGAAATAGATGAGACCTTCAGAGACT 300
Db 3289 AATATTGAACTGCACTCCCTGACTGCGAGAGAAATAGATGAGACCTTCAGAGACT 3348
Qy 301 CCAGGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGCGGCAAGCTGAGTGAT 360
Db 3349 CCAGGAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGCGGCAAGCTGAGTGAT 3408
Qy 361 CAAGGATCTGGAGCCGCTGGCGATCTCCTATTGACTCTCTCCAGATCACCTCGA 420
Db 3409 CAAGGATCTGGAGCCGCTGGCGATCTCCTATTGACTCTCTCCAGATCACCTCGA 3468
Qy 421 GAAAGTCAAGCACTTTCAGAGAGAAATTTGGCGCTCTGAAAGAGAAAGTGAAGCACCTCAA 480
Db 3469 GAAAGTCAAGCACTTTCAGAGAGAAATTTGGCGCTCTGAAAGAGAAAGTGAAGCACCTCAA 3528

Qy 481 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCAC 540
Db 3529 TGACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCAC 3588
Qy 541 TCTGAGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 600
Db 3589 TCTGAGAGACCTGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAG 3648
Qy 601 CGAGCTGCATGAAGCCCAACAGGAGCTTTTGTGTCAGCATCTCAGACATTTCTTTCCAGTTC 660
Db 3649 CGAGCTGCATGAAGCCCAACAGGAGCTTTTGTGTCAGCATCTCAGACATTTCTTTCCAGTTC 3708
Qy 661 TGTCCAGGTCCTTGGGAGAGCCATCTGCGCAAAACAAAGTGCCTACTATATCAACCA 720
Db 3709 TGTCCAGGTCCTTGGGAGAGCCATCTGCGCAAAACAAAGTGCCTACTATATCAACCA 3768
Qy 721 CGAGACTCAAAACAACTTGTGCGGACCATCCCAAAATGACAGAGCTTACCACTTTTACG 780
Db 3769 CGAGACTCAAAACAACTTGTGCGGACCATCCCAAAATGACAGAGCTTACCACTTTTACG 3828
Qy 781 TGACCTGAATTAATGTCAAGTTCCTCAGCTTATAGACTGCCATGAACTCCGAAAGCTGCA 840
Db 3829 TGACCTGAATTAATGTCAAGTTCCTCAGCTTATAGACTGCCATGAACTCCGAAAGCTGCA 3888
Qy 841 GAAGGCCCTTTGCTTGGATCTCTTGAGCTGTGATGCTGATGCTGATGCTGATGCTGATGCT 900
Db 3889 GAAGGCCCTTTGCTTGGATCTCTTGAGCTGTGATGCTGATGCTGATGCTGATGCTGATGCT 3948
Qy 901 CAAGCTCAAGCAAAATGACAGCCCATGGATATCTCAGATTAATTAATTTTGAACCA 960
Db 3949 CAAGCTCAAGCAAAATGACAGCCCATGGATATCTCAGATTAATTAATTTTGAACCA 4008
Qy 961 TATTATGACCGCTGGAGCAGAGCAACAATTTGGTCAAGTCTCTCTCGCTGGA 1020
Db 4009 TATTATGACCGCTGGAGCAGAGCAACAATTTGGTCAAGTCTCTCTCGCTGGA 4068
Qy 1021 TATGCTGCTGAATCTGCTGAATGTTTATGATACGGGACCAACAGAGAGTCCCGTGT 1080
Db 4069 TATGCTGCTGAATCTGCTGAATGTTTATGATACGGGACCAACAGAGAGTCCCGTGT 4128
Qy 1081 CCGTGTCTTTTAAACTGGCATCATTTCCCTGTGTGTAAGCACTTTGGAAGCAAGTACAG 1140
Db 4129 CCGTGTCTTTTAAACTGGCATCATTTCCCTGTGTGTAAGCACTTTGGAAGCAAGTACAG 4198
Qy 1141 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACCGAGCGAGGCTGGGCT 1200
Db 4189 ATACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACCGAGCGAGGCTGGGCT 4248
Qy 1201 CCTTCTGCATGATTTCTATCCAAATTCAGACAGTTCGGGTGAAGTTCATCTCTTTGGGG 1260
Db 4249 CCTTCTGCATGATTTCTATCCAAATTCAGACAGTTCGGGTGAAGTTCATCTCTTTGGGG 4308
Qy 1261 CAGTAACTTGAAGCAAGTTCGAGAGCTGCTTCCAAATTG 1301
Db 4309 CAGTAACTTGAAGCAAGTTCGAGAGCTGCTTCCAAATTG 4349

RESULT 14
US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR FILING DATE: 60/200,777
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36

LENGTH: 5060
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-36

Query Match 100.0%; Score 1301; DB 13; Length 5060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGATACAGGCGCTTCAAGAGGAATTGAA 60
DB 3119 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGATACAGGCGCTTCAAGAGGAATTGAA 3178

QY 61 AACTAAGAACCTGTATCATGATGACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
DB 3179 AACTAAGAACCTGTATCATGATGACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3238

QY 121 GCCTTTGGAGAGCTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCCTCGAGAGAG 180
DB 3239 GCCTTTGGAGAGCTAGAGAACTCTACAGAGAGCCAGAGAGCTGCCCTCGAGAGAG 3298

QY 181 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGAGGCTGAGAGGTCAATACAGTGGGA 240
DB 3299 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGAGGCTGAGAGGTCAATACAGTGGGA 3358

QY 241 AAAATTGAACCTGACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
DB 3359 AAAATTGAACCTGACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 3418

QY 301 CCAGGAACCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGGTGAT 360
DB 3419 CCAGGAACCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGGTGAT 3478

QY 361 CAGGGATTCCTGGAGCCCGCTGGGGGATCTCTCATTTGACTCTCTCCAGATCACTCGA 420
DB 3479 CAGGGATTCCTGGAGCCCGCTGGGGGATCTCTCATTTGACTCTCTCCAGATCACTCGA 3538

QY 421 GAAAGTCAAGGCATTCGAGAGGAAATTTGGCCCTCTGAAAGAGAACTGAGCAGCACTCAA 480
DB 3539 GAAAGTCAAGGCATTCGAGAGGAAATTTGGCCCTCTGAAAGAGAACTGAGCAGCACTCAA 3598

QY 481 TGACCTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTACCCGTATAAAGCTCAGCAC 540
DB 3599 TGACCTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTACCCGTATAAAGCTCAGCAC 3658

QY 541 TCTGGAAGACTGAAACACCAAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 600
DB 3659 TCTGGAAGACTGAAACACCAAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACCGAGTCAG 3718

QY 601 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTC 660
DB 3719 GCAGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGTC 3778

QY 661 TGTCAGGGTCCCTGGAGAGAGCCATCTGCCAACAAGTGGCCCTACTATATCAACCA 720
DB 3779 TGTCAGGGTCCCTGGAGAGAGCCATCTGCCAACAAGTGGCCCTACTATATCAACCA 3838

QY 721 CGAGCTCAAAACACTTTGCTGGGACCACTCCCAAAATGACAGAGCTCTACAGTCTTTTACG 780
DB 3839 CGAGCTCAAAACACTTTGCTGGGACCACTCCCAAAATGACAGAGCTCTACAGTCTTTTACG 3898

QY 781 TGACCTGAATTAATGTCAGATCTCAGCTTATAGCATGCGCATGAATCTCGAAGACTGCA 840
DB 3899 TGACCTGAATTAATGTCAGATCTCAGCTTATAGCATGCGCATGAATCTCGAAGACTGCA 3958

QY 841 GAAGGCCCTTTGCTTGATCTCTTGAGCCCTGTCAGCTGATGATGCTTGGACCAAGCA 900
DB 3959 GAAGGCCCTTTGCTTGATCTCTTGAGCCCTGTCAGCTGATGATGCTTGGACCAAGCA 4018

QY 901 CAACCTCAAGCAAAATGACCAAGCCCATGATCTCTGCAGATTAATTAATTTTGGACCA 960
DB 4019 CAACCTCAAGCAAAATGACCAAGCCCATGATCTCTGCAGATTAATTAATTTTGGACCA 4078

RESULT 15

US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIORITY FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIORITY FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27

Query Match 100.0%; Score 1301; DB 13; Length 5149;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGAATTGAA 60
DB 3208 CGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGAATTGAA 3267

QY 61 AACTAAGAACCTGTATCATGATGACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 120
DB 3268 AACTAAGAACCTGTATCATGATGACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 3327

QY 121 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCCTCTCAGGAGAG 180
DB 3328 GCCTTTGGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCCTCTCAGGAGAG 3387

QY 181 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGAGGTCAATACTGAGTGGGA 240
DB 3388 AGCCCAAGATGTCACTCGGCTTCTACGAAGCAGGCTGAGAGGTCAATACTGAGTGGGA 3447

QY 241 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 300
DB 3448 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACT 3507

QY 301 CCAGGAACCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 360
DB 3508 CCAGGAACCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGAT 3567

Qy 361 CAAGGGATCTTGGCAGCCCTGGCGGATCTCTCTCAATGACTCTCTCTCAAGATCACCTCCA 420
Db |||||
3568 CAAGGGATCTTGGCAGCCCTGGCGGATCTCTCTCAATGACTCTCTCTCAAGATCACCTCCA 3627
Qy 421 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCCCTCTGAAAGAGAGCTGAGCCACGTCAA 480
Db |||||
3628 GAAAGTCAAGGCACTTCGAGGAGAAATTCGGCCCTCTGAAAGAGAGCTGAGCCACGTCAA 3687
Qy 481 TGACCTTGGCTCGCCAGCTTACCACTTTGGCAATTCAGCTCTCACCGTATAAACCCTCAGCAC 540
Db |||||
3688 TGACCTTGGCTCGCCAGCTTACCACTTTGGCAATTCAGCTCTCACCGTATAAACCCTCAGCAC 3747
Qy 541 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCA 600
Db |||||
3748 TCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCAAGTGGCCGTGAGGACCGAGTCA 3807
Qy 601 GCAGCTCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 660
Db |||||
3808 GCAGCTCATGAAGCCACAGGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTC 3867
Qy 661 TGTCCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTCCCTACTATATCAACCA 720
Db |||||
3868 TGTCCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTCCCTACTATATCAACCA 3927
Qy 721 CGAGACTCAACAACTTGTGGGAGCCATCCCAATGACAGAGCTCTACAGTCTTTAGC 780
Db |||||
3928 CGAGACTCAACAACTTGTGGGAGCCATCCCAATGACAGAGCTCTACAGTCTTTAGC 3987
Qy 781 TGACCTGAATAATGTCAAGATTCAGCTTATAGGATGCCATGAAACTCCGAAGACTGCA 840
Db |||||
3988 TGACCTGAATAATGTCAAGATTCAGCTTATAGGATGCCATGAAACTCCGAAGACTGCA 4047
Qy 841 GAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCTTGGACAGCA 900
Db |||||
4048 GAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCAATGTGATGCTTGGACAGCA 4107
Qy 901 CAACCTCAAGCAAAATCAGCAGCCATGGATATCCTGCAGATTATTAATTTGTTGACCA 960
Db |||||
4108 CAACCTCAAGCAAAATCAGCAGCCATGGATATCCTGCAGATTATTAATTTGTTGACCA 4167
Qy 961 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA 1020
Db |||||
4168 TATTTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTGCGTGA 4227
Qy 1021 TATGCTGTGAGCTGCTGATGATGTTTATGATCGGAGCAAGAGGAGATCCGTGT 1080
Db |||||
4228 TATGCTGTGAGCTGCTGATGATGTTTATGATCGGAGCAAGAGGAGATCCGTGT 4287
Qy 1081 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 1140
Db |||||
4288 CCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAG 4347
Qy 1141 ATACCTTTTCAGCAAGTGGAGTTTCAACAGATTTTGTGACAGGCGAGGCTGGGCT 1200
Db |||||
4348 ATACCTTTTCAGCAAGTGGAGTTTCAACAGATTTTGTGACAGGCGAGGCTGGGCT 4407
Qy 1201 CCTTCTGATGATTTCTATCCAAATTCAGACAGTTGGGTGAAGTTGCATCCCTTTGGGG 1260
Db |||||
4408 CCTTCTGATGATTTCTATCCAAATTCAGACAGTTGGGTGAAGTTGCATCCCTTTGGGG 4467
Qy 1261 CAGTAACTGAGCAAGTGTCCGAGCTGCTTCCAAATTG 1301
Db |||||
4468 CAGTAACTGAGCAAGTGTCCGAGCTGCTTCCAAATTG 4508

Search completed: February 2, 2004, 17:38:55
Job time : 428.945 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:04:35 ; Search time 2439.51 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-10_COPY_1800_3100

Perfect score: 1301

Sequence: 1 cgaacttcagcagtgtaga.....ccggagctgctccaattg 1301

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_esthum:*

3: em_estnu:*

4: em_estov:*

5: em_estpl:*

6: em_estro:*

7: em_estc:*

8: gb_estl:*

9: gb_estl:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pla:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	584	44.9	824	9	AL556247
2	582.4	44.8	1098	13	BX365572
3	573	44.0	620	13	BQ640063
4	571.4	43.9	801	14	CB991394

ALIGNMENTS

RESULT 1
AL556247
LOCUS
DEFINITION AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CSODK001YB17 5-PRIME, mRNA sequence.
ACCESSION AL556247
VERSION AL556247.2 GI:31278051
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 824)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12898746.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation 1600
Faraday Avenue Genoscope sequence ID : CSODK001CA09QPI.
Location/Qualifiers

5	537.8	41.3	770	10	BG719710
6	535	41.1	3753	11	AK081426
7	526.4	40.5	1298	11	AK087829
8	526.4	40.5	2135	11	AK013510
9	526.4	40.5	4437	11	AK036936
10	524.8	40.3	1384	11	AK075809
11	511	39.3	797	14	CB960722
12	502.8	38.6	652	10	BB629984
13	491.6	37.8	704	10	BB610411
14	479.6	36.9	854	9	AI196693
15	470.8	36.2	717	14	CB527785
16	455.8	35.0	591	2	HS0083521
17	439	33.7	578	2	HS0075761
18	436.6	33.6	665	14	BY742804
19	431.4	33.2	898	10	BF182065
20	427.8	32.9	843	14	CA988247
21	426.6	32.8	3051	11	BC036095
22	410.4	31.5	599	10	BB666688
23	407.2	31.3	554	14	CB613696
24	404.4	31.1	493	14	CA888941
25	397.4	30.5	495	14	CA894775
26	392.2	30.1	650	14	BY714491
27	389	29.9	794	13	BU424348
28	385.2	29.6	684	9	AL641565
29	383.4	29.5	483	14	CA893902
30	381.8	29.3	423	9	AA460476
31	380.4	29.2	633	9	AL796733
32	375.4	28.9	514	9	AL871560
33	365.4	28.1	655	14	CA558919
34	349.6	26.9	546	4	EX516597
35	347.2	26.7	665	10	BE201973
36	337.6	25.9	541	9	AL894729
37	330	25.4	608	9	AL672616
38	329	25.3	532	10	BE334408
39	325.6	25.0	532	12	BU036615
40	325.6	25.0	883	14	CD557811
41	323.4	24.9	644	13	BU313510
42	320.6	24.6	406	14	CA871483
43	315.4	24.2	541	12	BI546771
44	313	24.1	656	9	AL869539
45	310	23.8	488	12	BM488542

C

AL556247 824 bp mRNA linear EST 31-MAY-2003
AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CSODK001YB17 5-PRIME, mRNA sequence.

ACCESSION AL556247
VERSION AL556247.2 GI:31278051

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 824)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT On Feb 15, 2001 this sequence version replaced gi:12898746.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 1955.r

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/InvitrogenCorporation 1600

Faraday Avenue Genoscope sequence ID : CSODK001CA09QPI.

Location/Qualifiers

FEATURES

source

```

source
1. .824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NciI-cligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
211 a 210 c 206 g 197 t
BASE COUNT
ORIGIN

```

RESIST. T. 2

EX365572	1098 bp	mRNA	linear	EST 05-MAY-2003
EX365572	1098 bp	mRNA	linear	EST 05-MAY-2003
LOCUS	EX365572	1098 bp	mRNA	linear
DEFINITION	EX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens			
	CDNA clone CS0DK001YB21 5-PRIME, mRNA sequence.			
ACCESSION	EX365572			
VERSION	EX365572.1	GI:30366927		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1098)			

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope

```
Query Match      44.8%; Score 582.4; DB 13; Length 1098;
Best Local Similarity 99.8%; Pred. No. 2.5e-159;
Matches 583; Conservative 0; Mismatches 1; Indels 0;
```


Library " 204 a 210 c 199 g 188 t

Query Match 43.9%; Score 571.4; DB 14; Length 801;
Best Local Similarity 99.7%; Pred. No. 3.5e-156;
Matches 583; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
718 CCACGAGACTCAACAACTTCTGGGACATCCCAAAATGACAGAGCTCTACAGTCTTT 777
Db 140 CCACGAGACTCAACAACTTCTGGGACATCCCAAAATGACAGAGCTCTACAGTCTTT 199
QY 778 AGCTGACCTGAATAATGTCAGATTCCTAGCTTATAGGACTGCCATGAACCTCCGAAGACT 837
Db 200 AGCTGACCTGAATAATGTCAGATTCCTAGCTTATAGGACTGCCATGAACCTCCGAAGACT 259
QY 838 GAGAGAGGCCCTTTGCTGGATCTCTGAGCTGTCAGCTGTCAGTGTGATGTCGCTGGACCA 897
Db 260 GAGAGAGGCCCTTTGCTGGATCTCTGAGCTGTCAGCTGTCAGTGTGATGTCGCTGGACCA 319
QY 898 GCACAACTCAAGCAAAATGACCCATGATATCTCGAGATTAATTAATGTTTGAC 957
Db 320 GCACAACTCAAGCAAAATGACCCATGATATCTCGAGATTAATTAATGTTTGAC 379
QY 958 CACTATTTATGACCCCTGGAGCAAGAGACAACTTTGTCAGCTCCCTCTCTGCGT 1017
Db 380 CACTATTTATGACCCCTGGAGCAAGAGACAACTTTGTCAGCTCCCTCTCTGCGT 439
QY 1018 GATATGTCCTGACCTGCTGCTGATGTTTATGATACGGGACCAAGAGAGATCCG 1077
Db 440 GATATGTCCTGACCTGCTGCTGATGTTTATGATACGGGACCAAGAGAGATCCG 499
QY 1078 TGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTA 1137
Db 500 TGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTA 559
QY 1138 CAGATACCTTTTCAAGCAAGTGCAGTTTCAACAGATTTTGTACCGACGCGCTGG 1197
Db 560 CAGATACCTTTTCAAGCAAGTGCAGTTTCAACAGATTTTGTACCGACGCGCTGG 619
QY 1198 CTCTCTTCTGCATGATCTTATCAAAATCCAAAGACAGTTGGTGAAGTTGATCCTTTGG 1257
Db 620 CTCTCTTCTGCATGATCTTATCAAAATCCAAAGACAGTTGGTGAAGTTGATCCTTTGG 679
QY 1258 GGCAGTAACATTGACCAAGTGTCC-GGAGCTGCTTCCAATTG 1301
Db 680 GGCAGTAACATTGACCAAGTGTCCGGAGCTGCTTCCAATTG 724

RESULT 5
BG719710
LOCUS 602690430F1 NIH_MGC_97 770 bp mRNA linear EST 08-MAY-2001
DEFINITION mRNA sequence.
ACCESSION BG719710
VERSION BG719710.1 GI:13998897
KEYWORDS Est.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 770)
NTH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10731 row: 1 column: 08
High quality sequence stop: 767.

FEATURES

source
1..770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4822807"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (grogag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 189 a 199 c 198 g 184 t

Query Match 41.3%; Score 537.8; DB 10; Length 770;
Best Local Similarity 97.4%; Pred. No. 2.5e-146;
Matches 568; Conservative 0; Mismatches 12; Indels 3; Gaps 2;
QY 718 CCACGAGACTCAACAACTTGTGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 777
Db 150 CCACGAGACTCAACAACTTGTGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 209
QY 778 AGCTGACCTGAATAATGTCAGATTCCTAGCTTATAGGACTGCCATGAACCTCCGAAGACT 837
Db 210 AGCTGACCTGAATAATGTCAGATTCCTAGCTTATAGGACTGCCATGAACCTCCGAAGACT 269
QY 838 GCAGAGAGGCCCTTTGCTGGATCTCTTGAGCTGTGAGCTGTGATGTCCTTGAGACCA 897
Db 270 GCAGAGAGGCCCTTTGCTGGATCTCTTGAGCTGTGAGCTGTGATGTCCTTGAGACCA 329
QY 898 GCACAA--CCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTAATTAATGTTTG 955
Db 330 GCACAACTCTCAAGTCAAGTACCAAGCCCATGGATATCTCGAGATTAATTAATGTTTG 389
QY 956 ACCACTATTATTATGACCGCTTGGAGCAAGACCAACAATTTGTTCAACCTCTCTCTGC 1015
Db 390 ACCACTATTATTATGACCGCTTGGAGCAAGACCAACAATTTGTTCAACCTCTCTGC 449
QY 1016 GTGGATATGTCGTGAATGCTGCTGAATGTTTATGATACGGGACCAAGAGGATC 1075
Db 450 GTGGATATGTCGTGAATGCTGCTGAATGTTTATGATACGGGACCAAGAGGATC 509
QY 1076 CGTGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAG 1135
Db 510 CGTGTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAG 569
QY 1136 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACGACGCGAGCTG 1195
Db 570 TACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACGACGCGAGCTG 628
QY 1196 GGCCTCTCTGCTGATGATTTCTCAAAATTCGAAGACAGTGGGTGAAGTTGATCCTTT 1255
Db 629 GGCCTCTCTGCTGATGATTTCTCAAAATTCGAAGACAGTGGGTGAAGTTGATCCTTT 688
QY 1256 GGGGGCAGTAACTGAGCAAGTGTCCGGAGCTGCTTCCAAT 1298
Db 689 GGGGGCAGTAACTGAGCAAGTGTCCGGAGCTGCTTCCAAT 731

RESULT 6
AK081426
LOCUS
DEFINITION Mus musculus 16 days embryo head cdna, RIKEN full-length enriched

AK081426 3753 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus 16 days embryo head cdna, RIKEN full-length enriched

library, clone:Cl30016K19 product:dystrophin related protein 2, full insert sequence.

ACCESSION AK081426 1 GI:26349154

VERSION AK081426 1 GI:26349154

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE 1
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20493374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format

JOURNAL sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakawa, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavola, H., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Koichiwa, H., Kueh, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Kuchikubush, J., Schraml, L. M., Stauber, P., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boiffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C., Claret, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, X. H., Weitz, C., Whitaker, C., Wilmig, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kotsuki, S., and Hayashizaki, Y.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation

JOURNAL Nature 409, 563-573 (2002)

REFERENCE 6

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hasizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE

JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES

source

1. 3753

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:Cl30016K19"

/db_xref="taxon:10090"

/clone="Cl30016K19"

/tissue_type="head"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="16 days embryo"

181..3054

/note="unnamed protein product; dystrophin related protein

2 [MGD|MGI:107432, GB|U43520, evidence: BLASTN, 99%,

match=807]

putative"

/codon_start=1

/protein_id="BAC38217.1"

/db_xref="GI:26349155"

/translation="MQPLVWQCGPYTLPRCHEWHAADRFHSSSLRNTCPQVRAAV

TIIPADWDGADPCLSFKLLNGVATGPLESAMLNWEIKKSHNLRLEAFSD

LSGKQLPRLIIDWSQKDELSAQLPGDVALVQKEKTHAAFEVKGKPYIS

SVLSQAQLFISHPFELBESQSDTPRQIRIQLSRFVWKQATVASLSEKFLAR

CVQDRIHETLLEHLEIOGMEELSLTQAGVRAWEIPGDLFIDSLPHIQAIK

LFKEEPSVKGDKVLYNDLAHOLAISDVHLSMENSRALEQINRKOLQVSAERLQK

LQDARDPFGSGSHFSTSVQVWEARFSPKVPYVNHQAQTCWDHPEKMTLYCTL

ADLNNIKFSAYRTAMKLRVQKALRDLDLTITLALIFNEHDLQAEHVMYVEVHC

LTALYERLEBERGILVNVPLVDMSLNLNLVDFDSGRGMRALSFKTGACLCGTEV

KEKLYLFSQVANGSQCORHLGALLHEAIVPQGLGEVAAFGVAVSPVSRCPFR

STGKPVIEASQFLWNLFPQSMWLAVLHRYVIAEQVKHQTCSICROCKINGFRYR

SLKQFNVDICQTCFLTRASKGNKLHPTMEYVPTTSSENVRDFAITLKNFRKQY

FSKHPORGLVQVSLSDSCSETPASSPMLPHADTHSRIEHPASRLAEMESQNSPFN

DSLSPDSIDEDQYLLRHSPITDRFPAQQAQPCSMATESKGELEKILAHLEENRI

LQELKRLQWHEEAAPTLVSGSAIPDRNEELLAEARILRHQKSRLETMRMIL

EDNKKQLSQRLRELLQPFSGDNGSAGSSLASPPQSGSHPREKNGQTTPUTE

VADVGSKSQVSLCEDINEKLHAFPSVRSSDVTANTLLAS"

3731..3736

/note="putative"

3753

/notes="putative"

/base_count 927 a 1021 c 938 g 867 t

polya_signal

polya_site

BASE COUNT 927 a 1021 c 938 g 867 t

ORIGIN

Query Match 41.1%; Score 535; DB 11; Length 3753;

Best Local Similarity 63.5%; Pred. No. 4.9e-145;

Matches 817; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

QY 14 GTTCAGAGGAGACGATGACATAGGCGCTTCAAGAGGGAATGAAACTAAGACCT 73

DB 592 GTACACAGGAGAGGAGACACATGACGCGCTTCAAGAGGGAATGAAACTAAGAGGCC 651

QY 74 GTAATCATGAGTACTCTTGGAGCTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGA 133
DB 652 TACATCTCTCTGTCTGTGAATCCGCCAGGCTTTCTGTCCCGAGCATCCATTTTGAAGAA 711
QY 134 CTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTC 193
DB 712 TTAGAGGAGTCTCAATCTGAGAGTAAGATACCTCCGCCAGACAGCGGATTTCAAACCTT 771
QY 194 ACTCGGCTTCTAGAAAGAGAGCTGAGGAGTCAATAGTGGGAAATTTGAACTTG 253
DB 772 AGTCGCTTTGTATGAGAGCAGCAACAGTGGCCAGTGTGAGCTGTGGGAAAGCTGACAGCC 831
QY 254 CACTCGCTGACTGGCAGAGAAATATAGATGAGACCTTTGAAAGACTCCAGGAATCTCAA 313
DB 832 CGCTGTGTAGATCAGATCGCCACATTTGAGATATCTCTGGAACATCTATTGGAGATCCAA 891
QY 314 GAGGCCACGATGAGCTGGAATCTCAAGCTGCGCCAGCTGAGGTGATCAAGGATCTCTGG 373
DB 892 GGGGCAATGGAGGAATTGAGCAGTACTTTTGACCCAGCAGAGGGAGTCCGAGCCACGCTGG 951
QY 374 CAGCCCGTGGCGATCTCTCTGATGACTCTCTCCAGATCACTCGAGAAAGTCAAGCA 433
DB 952 GAGCCATAGGAGATCTTTATTTGATTTCCCTCCAGAGCATATCCAGCCATCAAGCTA 1011
QY 434 CTTGAGGAGAAATTCGGCTCTGAAAGAGAACGTGAGCCAGCTCAATGACCTTGCTCGC 493
DB 1012 TTCAAAGAGAAATTTCTCTCTGTGAAAGATGGGTGAAGTGAATGATCTGCGCCAC 1071
QY 494 CAGCTTACACATTTGGCATTCAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 553
DB 1072 CAGCTTGCATCTCTGATGCACTTGTCAATGGAGATTCAGGGCTCTGGAACAGATC 1131
QY 554 AACACAGATGGAAGTCTTGAGGCTGGCGCTCGAGGACCGAGTCAAGCTGATGAA 613
DB 1132 AACATCGGTGGAACAGCTTCCAGGTGTGAGTGTGAGAGCTTAAAGCACTCCAGAT 1191
QY 614 GCCCAGAGGACTTGGTCAGATCTCAGACTTCTTTCAGCTCTCTCCAGGCTCC 673
DB 1192 GCCCAGCGGACTTGGCGCTGGGTGACAGCACTTCTCTCCACTTCTGTCAGAGTCTCC 1251
QY 674 TGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTATATCAACACAGAGACTCAACA 733
DB 1252 TGGGAAGAGCATTTCTCCATTAAGTCCCTACTATCAACACAGGCTCAGACC 1311
QY 734 ACTTGTGGAGCAATCCCAATATGACAGAGCTCTACAGCTTTTGTAGCTGACCTGATAT 793
DB 1312 ACATGCTGGGACCATCTAGAGTACTGAGTTATACCAACCTTAGCGGATCTGAACAC 1371
QY 794 GTCAGATTTCTCAGCTTATAGAGCTGCATGAACTCCGAGAGCTGAGAGGCTTTTGC 853
DB 1372 ATTAAGTTCTCAGCTTACCGCATCGCATGAGCTCGCAGAGTCCGAGAGGCTTGGC 1431
QY 854 TTGGATCTTTGAGCCTTGCAGCTGATGTGCTTCAAGCTCCCTCTCTGCGTGGATATGTCTG 913
DB 1432 CTGGATCTGTGTAATTTAACTACAGCTCTGAGATCTTCAATGAGCATGACTTGCAGGCC 1491
QY 914 ATGACAGAGCCATGATATCTTGCAGATTTAAATTTGTTGACCACTATTATGACGCG 973
DB 1492 AGTGAACATGTATGAGTGTGGTGGAGGTCATTCATCTGCTTGACTCCCTTGTATGAACA 1551
QY 974 CTGGAGCAAGAGCACAACAATTTGGTCAAGCTCCCTCTCTGCGTGGATATGTCTGAC 1033
DB 1552 CTGGAGAGGAGAGAGGATCTCTGTCATATGTGCCACTGTGTAGACATGAGCCTCAAC 1611
QY 1034 TGGTGTCTGAATTTATGATAGAGGAGCAAGAGGAGGATCCGTCTCTGCTGCTTTAA 1093
DB 1612 TGGCTCTCAATGTTTGTATGTGTGCTGAGTGGTGGAGGATGAGAGATGCTGCTTTAAG 1671
QY 1094 ACTGGCATCTTCCCTGTTGAAGACATTTGGAAGACAGTACAGATACCTTTTCAAG 1153
DB 1672 ACTGGCATCGATCTCTGTTGGCAGGAGTGAAGAAATCTTCAATCTCTTTCAGC 1731
QY 1154 CAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGAGGCTGGGCTCTCTGCTGATGAT 1213

DB 1732 CAAGTAGCCAAATTCAGGAGCCAGTGTGATCAACGCCATCTCGGTGCTCTTCATGAA 1791
QY 1214 TCTATCAAAATTCAGAGAGCTGGTGAAGTGTGATCTTTGGGGGAGTAAACATTTGAG 1273
DB 1792 GCATTCAGTCCCGCTCAGCTGGTGAAGTGGAGCATTTTGGGGGAGCAATGTGGAG 1851
QY 1274 CCAAGTGTCCGAGGCTGCTTCCAAATTT 1300
DB 1852 CCAGTGTCCGTAGCTGCTTTCGCTTT 1878
AK087829 1298 bp mRNA linear HTC 05-DEC-2002
Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN
full-length enriched library, clone:R330026B12 product:dystrophin,
muscular dystrophy, full insert sequence.
AK087829 1 GI:26352777
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493374
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Fujiwara, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Yamamoto, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Tonita, M.,
Quackenbush, J., Schriml, L.M., Staib, F., Suzuki, R., Aono, H.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bernaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection

DB 469 CAGATACCTTTTCAAGCAGTGGGCAAGTTCACCTGGCTTTTGTGACACGCTAGGCTGGG 528
 QY 1198 CTTCTTCTTCGATGATTTCTATCCAAATTCGAAGAGTGGTGAAGTTCATCTTTGG 1257
 DB 529 TCTTCTTCTGATGATTTCTATTCAAATCCCAAGCAGTGGTGAAGTTCCTTCTTTGG 588
 QY 1258 GGGCAGTACATGAGCCAGTGTCCGAGTGTCTTCCAAATTTG 1301
 DB 589 GGGCAGTAACTGAGCCAGTGTCCGAGTGTCTTCCAAATTTG 632

RESULT 10
 AK075809
 LOCUS
 DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
 enriched library, clone:1810074E05 product:dystrophin, muscular
 dystrophy, full insert sequence.
 ACCESSION AK075809
 VERSION 1 GI:26344655
 KEYWORDS HTC; CAP trapper;
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishie, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gocho, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kouchiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nika, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barah, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Nordone, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohtsuki, S.,
 Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 6 (bases 1 to 1384)
 PUBMED 12053913
 REFERENCE 6
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Azai, A.,
 Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
 Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
 Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
 Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
 Saito, H., Saito, R., Sakai, K., Sano, H., Sasaki, D.,
 Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
 Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
 Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
 Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in RIKEN
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 FEATURES
 source
 1. 1384
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:1810074E05"
 /db_xref="MGI:1908112"
 /db_xref="taxon:10090"
 /clone="1810074E05"
 /sex="male"
 /tissue_type="pancreas"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="10 day old"
 120_1142
 /note="unnamed protein product; dystrophin, muscular
 dystrophy (MGI:194909, GB|NM_007868, evidence: BLASTN,
 100%, match=999)
 putative"
 /codon_start=1
 /protein_id="BAC35977.1"
 /db_xref="GI:26344656"
 /db_xref="MGI:194909"
 /translation="MEHLKGHTQTQWDHPKWTLYOSLADLNNVRSAYRTAKML
 RRLQKALDLSLSAACDADQNLKQNDQPMDLQINCLTYDLEQEHNNLVN
 VPLCVDCNLNLLNVDYDGTGRIVLSFTKSLKARLEKYLKQVASSGTG
 CDQRLLGLLHDSIQIPLQGEVASFSGNIEPSVSCFOFANNKPEIEALFLDNR
 LEPOSVMVLPVLRVAATAKQCNICKECPIIGFRRYSKHFNYDQCSCFFSG
 RVAKGHKVMYWEYCTPTTSGEDVDFAKVLKNKFKTRFKYFAKHPRMGVLPQTILE
 GDNMET"
 BASE COUNT 376 a 354 c 303 g 351 t
 ORIGIN
 Query Match 40.3%; Score 524.8; DB 11; Length 1384;
 Best Local Similarity 93.7%; Pred. No. 2.5e-142;
 Matches 547; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY	718	CCACGAGCTCAAAACAACCTTGCTGGGACCATCCAAAATGACAGAGCTCTACCAAGTCTTT	777
Db	140	CCACGAGACCAAAACAACCTTGTTGGGACCAACCACAAAATGACAGAGCTCTACCAAGTCTTT	199
QY	778	AGCTGACCTGAATAATGTCTAGATTCTTCAGCTTATAGGACCTGCCATGAAACTCCGAGAGCT	837
Db	200	AGTGTACCTGAAATAATGTCAAGTTCTCCGCGTATAGGACTTGCATGAAGCTCAGAAGGCT	259
QY	838	GCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAAGCTGTCAATGTGATGCCCTTGACCA	897
Db	260	CCAGAAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCCGCTGCAATGTATGCCCTTGACCA	319
QY	898	GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTCGAGATTATTAATGTTTGAC	957
Db	320	GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTCGAGATTATTAATGTTTGAC	379
QY	958	CACATATTTATGACCGCTGTGAGCAAGAGCACAAATTTGGTCAAAGCTCCCTCTCTGCGT	1017
Db	380	TACAAATTTATGATCGTCTGGAGCAAGAGCACAAATCGGTCAATGTCCTCTCTGTGT	439
QY	1018	GGATATGTGTCTGAACTGGCTGCTGTAATGTTTATGATACGGAGCAACAGGAGGATCCG	1077
Db	440	GGATATGTGTCTCAACTGGGCTTCTCAATGTTTATGATACGGAGCAACAGGAGGATCCG	499
QY	1078	TGTCCTGTCTTTTAAAACTGGCATCTTTCCCTGTGTAAAGACATTTGGAAGACAAGTA	1137
Db	500	TGTCCTGTCTTTTAAAACTGGCATCTTTCTGTGTAAAGACACATTTGGAAGACAAGTA	559
QY	1138	CAGATACCTTTTCAAGCAAGTGGCAAGTTCAAACAGATTTTGTGTACAGCGCAGGCTGGG	1197
Db	560	CAGATACCTTTTCAAGCAAGTGGCAAGTTCAAACAGGCTTTTGTACAGCGCTAGGCTGGG	619
QY	1198	CCCTCCTCTGCATGATTCATCCAAATTCAGACAGTTGGGTGAAGTTGCATCCCTTTGG	1257
Db	620	TCCTCTCTGCATGATTCATTCAAATCCCAAGACAGTTGGGTGAAGTTGCTTCCTTTGG	679
QY	1258	GGGCAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG	1301
Db	680	GGGCAGTAACATTGAGCCAAAGTGTCCGGAGCTGCTTCCAAATTTG	723

```

RESULT 11
CB960722
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CB960722 797 bp mRNA linear EST 29-APR-2003
AGENCOUT 13761995 NIH_MGC 147 Homo sapiens cDNA clone
IMAGE:30344479 5', mRNA sequence.
CB960722
CB960722.1 GI:30216839
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 797)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Dr. Stefan Hanson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E.B. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM380 row: n. column: 08
High quality sequence stop: 620.
Location/Qualifiers
1..797
/organism="Homo sapiens"
FEATURES
source

```

```

/mol_type="nrna"
/db_xref="taxon:9606"
/clone="IMAGE:3034479"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 147"
/note="Organ: placenta; Vector: pBluescriptR; Site 1:
ali-XhoI; Site 2: BamI; Oligo-dr primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."
BASE COUNT      199 a      208 c      199 g      191 t
ORIGIN
      Query Match      39.3%; Score 511; DB 14; Length 797;
      Best Local Similarity 99.8%; Pred. No. 1.9e-138;
      Matches 522; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 718 CCACGAGACTCAAAACACTTGTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 777
Db 168 CCACGAGACTCAAAACACTTGTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 227

QY 778 AGCTGACCTGAAATAAGTCAGATCTCAGCTTATAGGACTGCCATCAAACTCCGAGAGACT 837
Db 228 AGCTGACCTGAAATAAGTCAGATCTCAGCTTATAGGACTGCCATCAAACTCCGAGAGACT 287

QY 838 GCAGAAGGCCCTTTGGTCTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCTTGACCA 897
Db 288 GCAGAAGGCCCTTTGGTCTGGATCTCTTGAGCCTGTGCAGCTGCATGTGATGCCTTGACCA 347

QY 898 GCACACCTCAAGCAAAATGACAGCCCATGGATATCTTCAGATTAATTAATTTGTTGAC 957
Db 348 GCACAACTCAAGCAAAATGACAGCCCATGGATATCTTCAGATTAATTAATTTGTTGAC 407

QY 958 CACTATTATGACCGCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGT 1017
Db 408 CACTATTATGACCGCGCTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGT 467

QY 1018 GGATATGTCTGAACCTGGCTGCTGTAATTTTATGATACGGAGCAACAGGAGGATCCG 1077
Db 468 GGATATGTCTGAACCTGGCTGCTGTAATTTTATGATACGGAGCAACAGGAGGATCCG 527

QY 1078 TGTCTCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGACACATTTGGAAGCAAGTA 1137
Db 528 TGTCTCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGACACATTTGGAAGCAAGTA 587

QY 1138 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACAGCGCAGGCTGGG 1197
Db 588 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACAGCGCAGGCTGGG 647

QY 1198 -CTCTCTCTGCGATGTTCTATCCAAATTCAGACAGATTTGGG 1239
Db 648 CCTCTCTCTGCGATGTTCTATCCAAATTCAGACAGATTTGGG 690

RESULT 12
BB629984
LOCUS BB629984 652 bp mRNA linear EST 31-AUG-2001
DEFINITION musculus cDNA clone 930028B14 5', mRNA sequence.
ACCESSION BB629984
VERSION BB629984.1
KEYWORDS GI:15399733
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 652)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.

```


Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yananaka, I., Aizawa
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues. Location/Qualifiers
 1. 704
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="1810074E05"
 /sex="male"
 /tissue types="pancreas"
 /dev stage="10 day old"
 /clone lib="RIKEN full-length enriched, 10 day old male
 pancreas"
 BASE COUNT 173 a 185 c 167 g 179 t
 ORIGIN
 Query Match 37.8%; Score 491.6; DB 10; Length 704;
 Best Local Similarity 93.0%; Pred. No. 8.8e+133;
 Matches 515; Conservative 0; Mismatches 39; Indels 0; Gaps 0
 Qy 718 CCAGGAGCTCAACAACCTTGCTGGGACCATCCCAAAATGACAGAGCTCTTACAGTCTTT 777
 Db 137 CCACGAGACCCAAACCACTTGTTGGGACCAACCCCAAAATGACAGAGCTCTTACAGTCTTT 196
 Qy 778 AGCTGAGCTGAATAATGTCAGATTCTCAGCTTATAGGAGTCCCATGAACCTCGAAGACT 837
 Db 197 AGTGAGCTGAATAATGTCAGGTTCTCCGCGTATAGGAGTCCCATGAAGCTCAGAAGCT 256
 Qy 838 GCAGAGGCCCTTTGCTTGATCTCTTGAGCTGTGAGCTGTGATGCTGATGCTGATGCTGACCA 897
 Db 257 CCAGAGGCCCTTTGCTTGATCTCTTGAGCTGTGAGCTGTGATGCTGATGCTGATGCTGACCA 316
 Qy 898 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTCGAGATTATTAATGTTTGAC 957
 Db 317 GCACAACTCAAGCAAAATGACAGCCCATGGATATCCTCGAGATTATTAATGTTTGAC 376
 Qy 958 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCTCTCTGCGT 1017
 Db 377 TACAATTTATGATCGCTGGAGCAAGAGCAACAATCTGGTCAATGTCCTCTCTGCTGT 436
 Qy 1018 GGATATGCTGTAACCTGGCTGCTGAATGTTATGATACGGAGCGACACGGAGGATCCG 1077
 Db 437 GGATATGCTGCTCAACTGGCTTCTCAATGTTATGATACGGAGCGACACGGAGGATCCG 496
 Qy 1078 TGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACATTTGGAAGACAGTA 1137

This clone was contributed by the Brain Molecular Anatomy Project (BMAP).

Seq Primer: pYX-5.

FEATURES
source
Location/Qualifiers
1. 717
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6850599"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DHI0B (Ti phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/notes="Organ: Brain; Vector: pYX-Asc; Site: Ecor I; Site 2: Not I; The library was constructed according to Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag is AGGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System, supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 180 a 167 c 171 g 198 t 1 others
ORIGIN

Query Match 36.2%; Score 470.8; DB 14; Length 717;
Best Local Similarity 93.9%; Pred. No. 1.2e-126;
Matches 490; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

780 CTGACCTGTAATATGTGATCTCAGCTTATAGGACTGCCATGAACTCCGAGACTGC 839
Db 1 CTGACCTGTAATATGTCAGGTTCTCCGGTATAGGACTGCCATGAACTCAGAAGCTCC 60

840 AGAAGGCCCTTTGGTTGGATCTCTTGAGCCTGTGAGTGTGATGCTTGACGACG 899
Db 61 AGAAGGCCCTTTGGTTGGATCTCTTGAGCCTGTGAGTGTGATGCTTGACGACG 120

900 ACAACCTCAAGCAAAATGACCGCCATGATATCTGCGAGATTATTAATTGTTGACCA 959
Db 121 ACAACCTCAAGCAAAATGACCGCCATGATATCTGCGAGATTATTAATTGTTGACCA 180

960 CTATTTATGACCGCTCGAGCAAGACACAAATTTGGTCAACGTCCTCTCTCTGCTGG 1019
Db 181 CAATTTATGATCGTCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTCTGCTGG 240

1020 ATATGCTCTGAAGTGGCTCTGATGTTTATGATACGGGACCAAGGAGGATCCGTG 1079
Db 241 ATATGCTCTGAAGTGGCTCTGATGTTTATGATACGGGACCAAGGAGGATCCGTG 300

1080 TCCCTGCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACA 1139
Db 301 TCCCTGCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGAAGACAAGTACA 360

1140 GATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTGTTGACCGGAGGAGGATCCGCT 1199
Db 361 GATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTGTTGACCGGAGGAGGATCCGCT 420

1200 TCCCTGCTGATGATTTCTATCCAAATTCAGACAGATTGGGTGAAGTTGCAATCTTTGGG 1259
Db 421 TCCCTGCTGATGATTTCTATCCAAATTCAGACAGATTGGGTGAAGTTGCAATCTTTGGG 480

1260 GCAGTAACATTGACCGAAGTGTCCGAGCTGCTTCCAAATTG 1301
Db 481 GCAGTAACATTGACCGAAGTGTCCGAGCTGCTTCCAAATTG 522

BASE COUNT 202 a 232 c 208 g 209 t 3 others
ORIGIN

Query Match 36.9%; Score 479.6; DB 9; Length 854;
Best Local Similarity 90.9%; Pred. No. 3.4e-129;
Matches 531; Conservative 0; Mismatches 51; Indels 2; Gaps 2;

718 CCACGAGACTCAACAACTGCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTT 777
Db 135 CCACGAGACCAACCACTGTTGGGACCAACCCCAAAATGACAGAGCTCTACCACTCTTT 194

778 AGCTGACCTCAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAGACT 837
Db 195 AGCTGACCTCAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAGACT 254

838 GCAGAGGCCCCCTTGGTTGGATCTCTGAGCCTGTCAGCTGATGATGATGCTTGGACCA 897
Db 255 CCAGAGGCCCCCTTGGTTGGATCTCTGAGCCTGTCAGCTGATGATGATGCTTGGACCA 314

898 GCACAACTCAAGCAAAATGACCAAGCCCATGGATATCTCGCAGATTATTAAATTGTTGAC 957
Db 315 GCACAACTCAAGCAAAATGACCAAGCCCATGGATATCTCGCAGATTATTAAATTGTTGAC 374

958 CACTATTTATGACCGCTGGAGCAAGACACAAATTTGGTCAAGTCCCTCTCTGCGT 1017
Db 375 TACAATTTATGATCGTCTGGAGCAAGACACAAATCTGGTCAATGTCCTCTCTGCTGT 434

1018 GGATATGTCCTCAAGTGGCTGCTGATGTTATGATACGGGACCAAGGAGGATCC- 1076
Db 435 GGATATGTCCTCAAGTGGCTTCTCAAGTTTATGATACGGGACCAAGGAGGATCCG 494

1077 GTGTCCTGCTTTTAAACCTGGCATCTTCCCTGTGTTAAAGCACATTTGGAAGCAAGT 1136
Db 495 GGTGCTGCTTTTAAACCTGGCATCTTCTGTGTTAAAGCACACTTTGGAAGCAAGT 554

1137 ACAGTACCTTTTCAAGCAAGTGGAGTCAACAGATTGTTGACGAGCCAGCGCTGG 1196
Db 555 ACAGTACCTTTTCAAGCAAGTGGAGTCAACAGATTGTTGACGAGCCAGCGCTGG 614

1197 GCTCTCTCTGCTGATGTTCTATCCAAATTCAGACAGTTGGGTGAAGTTGATCCTTTG 1256
Db 615 GTCTCTCTCTGCTGATGTTCTATCCATCCCAAGACAGTGGTGAAGTGTCTTCTTT- 673

1257 GGGGAGTAACTGACCAAGTGTCCGAGCTGCTTCCAAATT 1300
Db 674 GGGGAGTAACTGACCAAGTGTCCGAGCTGCTTCCAAATT 717

RESULT 15
CB527785
LOCUS 717 bp mRNA linear EST 28-MAR-2003
DEFINITION UI-M-FY0-cfk-c-14-0-UI.r1 NIH BMAP_FY0 Mus musculus cDNA clone
INAME: 6850599 5', mRNA sequence.
ACCESSION CB527785
VERSION 1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 717)
NTH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>

Search completed: February 2, 2004, 06:03:15
Job time : 2442.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:14:55 ; Search time 7484.44 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-12_COPY_900_3000

Perfect score: 2101

Sequence: 1 gagctatgctacacacagg.....ttctgatgattctatccaa 2101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:**

1: gb_ba:**

2: gb_btg:**

3: gb_in:**

4: gb_om:**

5: gb_ov:**

6: gb_pat:**

7: gb_ph:**

8: gb_pl:**

9: gb_pr:**

10: gb_ro:**

11: gb_sta:**

12: gb_sy:**

13: gb_un:**

14: gb_vi:**

15: em_ba:**

16: em_fun:**

17: em_hum:**

18: em_in:**

19: em_mu:**

20: em_om:**

21: em_or:**

22: em_ov:**

23: em_pat:**

24: em_ph:**

25: em_pl:**

26: em_ro:**

27: em_sta:**

28: em_sy:**

29: em_un:**

30: em_vt:**

31: em_htg_hum:**

32: em_htg_inv:**

33: em_htg_other:**

34: em_htg_mu:**

35: em_htg_pln:**

36: em_htg_rod:**

37: em_htg_man:**

38: em_htg_vrt:**

39: em_sy:**

40: em_htgo_hum:**

41: em_htgo_mu:**

42: em_htgo_other:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1760.4	83.8	5339	6	AX538620	AX538620 Sequence
2	1629	77.5	5462	6	AX538621	AX538621 Sequence
3	1611.6	76.7	5417	6	AX538619	AX538619 Sequence
4	1328.8	63.2	5952	6	AR304538	AR304538 Sequence
5	1328.8	63.2	5952	6	AX114289	AX114289 Sequence
6	1328.8	63.2	8689	6	AX538622	AX538622 Sequence
7	1328.8	63.2	11443	6	AX538624	AX538624 Sequence
8	1328.8	63.2	12057	6	AX538627	AX538627 Sequence
9	1328.8	63.2	12446	9	HSMDMR	X14298 Human mRNA
10	1328.8	63.2	13957	6	AX409637	AX409637 Sequence
11	1328.8	63.2	13957	6	AX538581	AX538581 Sequence
12	1328.8	63.2	13957	9	HUMDYS	M18533 Homo sapien
13	1317.8	62.7	13977	6	AR220819	AR220819 Sequence
14	1317	62.7	4402	6	E30219	E30219 Shortened d
15	1196.6	57.0	4402	6	E30220	E30220 Shortened d
16	1185.4	56.4	13887	4	AF070485	AF070485 Canis fam
17	1149.6	54.7	13815	6	AX306153	AX306153 Sequence
18	1149.6	54.7	13815	6	AX538582	AX538582 Sequence
19	1148.6	54.7	13815	10	MUSDYSA	M68859 Mouse dystr
20	1149.6	54.7	19307	6	AR093392	AR093392 Sequence
21	1149.6	54.7	19307	6	AR142592	AR142592 Sequence
22	922.2	43.9	13575	5	GGDYS	X13369 Chicken mRN
23	718.8	34.2	4075	6	E30221	E30221 Shortened d
24	718.8	34.1	2654	5	FSCDYSTRO	M37645 Torpedo cal
25	667.2	31.8	3747	6	E30218	E30218 Shortened d
26	623.8	30.0	3275	10	MUSDYS	M18025 Mouse dystr
27	613.6	29.2	3163	6	E30223	E30223 Shortened d
28	598.2	28.5	3521	5	AF339031	AF339031 Danio rer
29	555.6	26.4	110705	10	RNAJ2967	AJ002967 Rattus no
30	554.2	26.4	11096	6	AX538584	AX538584 Sequence
31	554.2	26.4	11096	10	MMY12229	Y12229 M.musculus
32	554	26.4	3161	10	MMGUTRPH	X83506 M.musculus
33	547.4	26.1	6045	6	A63605	A63605 Sequence 7
34	547.4	26.1	6045	6	AR281528	AR281528 Sequence
35	547.4	26.1	6059	6	AX107972	AX107972 Sequence
36	547.4	26.1	10302	9	AX538583	AX538583 Sequence
37	547.4	26.1	10302	9	HSMUPS	X69086 H.sapiens m
38	547.4	26.1	10320	6	A63607	A63607 Sequence 9
39	547.4	26.1	10320	6	AR281529	AR281529 Sequence
40	528.4	25.1	3073	10	AF195788	AF195788 Rattus no
41	528.4	25.1	3140	10	AF195787	AF195787 Rattus no
42	527.4	25.1	3499	9	HSU43519	U43519 Human dystr
43	527.4	25.1	5106	6	AX552248	AX552248 Sequence
44	505	24.0	2110	9	HUMDMDX	M92650 Human Duch
45	505	24.0	4658	9	BC028720	BC028720 Homo sapi

ALIGNMENTS

RESULT 1	AX538620	5339 bp	DNA	linear	PAT 23-NOV-2002
AX538620	Sequence 40 from Patent WO0229056.				
LOCUS	AX538620				
DEFINITION	Sequence 40 from Patent WO0229056.				
ACCESSION	AX538620				
VERSION	AX538620.1	GI:25271166			
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
REFERENCE	1				
AUTHORS	Chamberlain, J.S. and Harper, S.Q.				
TITLE	Mini-dystrophin nucleic acid and peptide sequences				
JOURNAL	Patent: WO 0229056-A 40 11-APR-2002;				
	THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)				

FEATURES		Location/Qualifiers	
source	1..5339	/organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630" /note="Synthetic"	
BASE COUNT	1638 a	1191 c	1187 g 1323 t
ORIGIN			
Query Match	83.8%;	Score 1760.4;	DB 6; Length 5339;
Best Local Similarity	90.3%;	Pred. No. 0;	
Matches 1910;	Conservative	0; Mismatches 191;	Indels 15; Gaps 2;
QY	1	GAGCTATGCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGAGGCCATT	60
Db	1099	GAGCTATGCTACACACAGGCTGCTTATGTCACCACTCTGACCCCTACACGAGGCCATT	1158
QY	61	TCCTTCACAGCATTTGGAGCTCTGGAAGCTCTGGAAGCAAGTCATTGTCAGTTCAATGATGGAGAG	120
Db	1159	TCCTTCACAGCATTTGGAGCTCTGGAAGCTCTGGAAGCAAGTCATTGTCAGTTCAATGATGGAGAG	1218
QY	121	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAAGATGATTAATGTCGTCCTTCTTTTC	180
Db	1219	TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAAGATGATTAATGTCGTCCTTCTTTTC	1278
QY	181	TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	240
Db	1279	TGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA	1338
QY	241	CCAGTTTCATACATCAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTTGG	300
Db	1339	CCAGTTTCATACATCAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTTGG	1398
QY	301	TAATATTTCTACAATTGGGAAGTAAAGCTGATGGAACAGGAAATATATCAGAAGATGAAGA	360
Db	1399	TAATATTTCTACAATTGGGAAGTAAAGCTGATGGAACAGGAAATATATCAGAAGATGAAGA	1458
QY	361	AACGTAAGTACAGAGAGATGATCTCTTAATTCAGATGGGATGCTCAGGGTAGC	420
Db	1459	AACGTAAGTACAGAGAGATGATCTCTTAATTCAGATGGGATGCTCAGGGTAGC	1518
QY	421	TAGCATGGAATAAACAAGCAATTTACATAGATGTTTAAATGGATCTCCAGAATCAGAACT	480
Db	1519	TAGCATGGAATAAACAAGCAATTTACATAGATGTTTAAATGGATCTCCCCAAGCT	1578
QY	481	GAAGAGTTGATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAATGAGGAGAA	540
Db	1579	GAAGAGTTGATGACTGGCTAACAAAAACAGAAAGAAAGAACAGGAAATGAGGAGATC	1638
QY	541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAAGGTGCTTCA	600
Db	1639	TACCCGTAAGGAAGGCTCTTAGAGACTCCAGGGGATGAAGAGCTGATGAACAATG	1698
QY	601	AGAAAGATCTAGAACAGAACAGTACGGGTCAATTTCTCACTCAATGGTGGTGTAGT	660
Db	1699	GCAAGACCTCCAAAGTGAAATTTGAAGCTCACACAGATGTTTATCACAACTGGATGAAA	1758
QY	561	TGATGA-----ATCTAGTGGAGATCAGCACTGCTGCTTTGGAGAACCACT	708
Db	1759	CAGCCAAAATCTAGATCCCTGGAGGTTCCGATGATGCGAGTCTCTGTTTACAAAGACG	1818
QY	709	TAAGG---TATTTGGAGATCGATGGGCAACATCTGTAGATGGAACAGACCGCTGGGT	765
Db	1819	TTTGGATAACATGAATTCAGTGGAGTGAATTCGGGAAAAGTCTCTCAACATTAGGTC	1878
QY	766	TCTTTTACAGACAGTTCTGACAGTGGAGCGTCTGCACCTTCTCTGCGAGAACTTCT	825
Db	1879	CCATTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGCACCTTCTCTGCGAGAACTTCT	1938
QY	826	GGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGACCTATTGGAGCGCACTT	885
Db	1939	GGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGACCTATTGGAGCGCACTT	1998
QY	886	TCACAGCAGTTTCAAGACGACGATGTATCATAGGCGCTTCAAGAGGGAATTCGAAACTAA	945
Db	1999	TCACAGCAGTTTCAAGACGACGATGTATCATAGGCGCTTCAAGAGGGAATTCGAAACTAA	2058
QY	946	AGAACTGTAAATCAATGAGTACTCTTTGAGACTCTAGCAATATTTCTGACAGAGCAGCTTT	1005
Db	2059	AGAACTGTAAATCAATGAGTACTCTTTGAGACTCTAGCAATATTTCTGACAGAGCAGCTTT	2118
QY	1006	GGAAGGACTAGAGAAACTCTTACCAGAGGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCA	1065
Db	2119	GGAAGGACTAGAGAAACTCTTACCAGAGGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCA	2178
QY	1066	GAATGTCACTCGGCTCTTACGAAAGCAGCTGAGNGTCAATCTGAGTGGGAAAAAT	1125
Db	2179	GAATGTCACTCGGCTCTTACGAAAGCAGCTGAGNGTCAATCTGAGTGGGAAAAAT	2238
QY	1126	GAACTCTGCATCTCGCTGACTTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGA	1185
Db	2239	GAACTCTGCATCTCGCTGACTTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGA	2298
QY	1186	ACTTCAAGAGGCCACGAGTGAAGCTCAAGCTGCGCCAGCTGAGGTGATCAAGGG	1245
Db	2299	ACTTCAAGAGGCCACGAGTGAAGCTCAAGCTGCGCCAGCTGAGGTGATCAAGGG	2358
QY	1246	ATCTGTCGACCGCTGGCGGATCTCTCTATTGACTCTCTCCAAGATCACTCCGAGAAAGT	1305
Db	2359	ATCTGTCGACCGCTGGCGGATCTCTCTATTGACTCTCTCCAAGATCACTCCGAGAAAGT	2418
QY	1306	CAAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAAAGCTGAGCCACGTCAATGACCT	1365
Db	2419	CAAGGCACTTCGAGGAGAAATTCGGCTCTGAAAGAGAAAGCTGAGCCACGTCAATGACCT	2478
QY	1366	TGCTCGCCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGTATAAAGCTCAGCACTCTGGA	1425
Db	2479	TGCTCGCCAGCTTACCACTTTGGGCACTTCAGCTCTCACCGTATAAAGCTCAGCACTCTGGA	2538
QY	1426	AGACCTGAAACACAGATGGAAGCTTTCGAGGTGGCGCTGAGGAGCCGAGTCAAGGAGCT	1485
Db	2539	AGACCTGAAACACAGATGGAAGCTTTCGAGGTGGCGCTGAGGAGCCGAGTCAAGGAGCT	2598
QY	1486	GCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTGTGCCA	1545
Db	2599	GCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTGTGCCA	2658
QY	1546	GGGTCCCTGGGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACACAGAGAC	1605
Db	2659	GGGTCCCTGGGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACACAGAGAC	2718
QY	1606	TCAAAACAACTTGTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGTGAACCT	1665
Db	2719	TCAAAACAACTTGTGGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGTGAACCT	2778
QY	1666	GAATAATGTCAAGTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCGAGAAGGC	1725
Db	2779	GAATAATGTCAAGTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCGAGAAGGC	2838
QY	1726	CCTTTGCTTGGATCTCTTGAGCCCTGTGAGTGCATGTGATGCTTGGACGACCAACCT	1785
Db	2839	CCTTTGCTTGGATCTCTTGAGCCCTGTGAGTGCATGTGATGCTTGGACGACCAACCT	2898
QY	1786	CAAGCAAAATGACCCAGCCATGGATATCTGCAATTAATTAATTTGTTGACCACTATTTA	1845
Db	2899	CAAGCAAAATGACCCAGCCATGGATATCTGCAATTAATTAATTTGTTGACCACTATTTA	2958
QY	1846	TCACCCCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGATATGTG	1905
Db	2959	TGACCCCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGATATGTG	3018
QY	1906	TCTGAACCTGGCTGTGTAATTTATGATACGGGAGCAAGAGGAGGATCCGCTGCTCTGC	1965
Db	3019	TCTGAACCTGGCTGTGTAATTTATGATACGGGAGCAAGAGGAGGATCCGCTGCTCTGC	3078
QY	1966	TTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTTGAAGACAAAGTACAGATACCT	2025

Db 3079 TTTTAAACATGGCATCTTCCCTGTGTAAAGCACATTTGGAAAGACAAGTACAGATACCT 3138
QY 2026 TTTCAAGCAAGTGGCAAGTTCAACAGATTTTCTGACCGAGCGAGCTGGCGCTCTTCT 2085
Db 3139 TTTCAAGCAAGTGGCAAGTTCAACAGATTTTCTGACCGAGCGAGCTGGCGCTCTTCT 3198
QY 2086 GCATGATTTATCCAA 2101
Db 3199 GCATGATTTATCCAA 3214

RESULT 2
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dyserophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 41 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES
Location/Qualifiers
source 1..5462
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"

BASE COUNT 1668 a 1225 c 1212 g 1357 t
ORIGIN
Query Match 77.5%; Score 1629; DB 6; Length 5462;
Best Local Similarity 85.4%; Pred. No. 0;
Matches 1911; Conservative 0; Mismatches 190; Indels 138; Gaps 3;

QY 1 GAGCTATGCCCTACACACAGCGCTGTATGTCACACCTCTGACCCCTACACGGAGCCCAT 60
Db 1099 GAGCTATGCCCTACACACAGCGCTGTATGTCACACCTCTGACCCCTACACGGAGCCCAT 1158
QY 61 TCCTTCACAGATTTGGAAGCTCTGAAGCAAGTCAATTTGGCAGTTCAATGATGAGAG 120
Db 1159 TCCTTCACAGATTTGGAAGCTCTGAAGCAAGTCAATTTGGCAGTTCAATGATGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAACAGCGTTTAGAAGAGTATTATCGTGGCTCTTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTATCAACAGCGTTTAGAAGAGTATTATCGTGGCTCTTTTC 1278
QY 181 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGGAGTGGTGAAGA 240
Db 1279 TGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAATGATGGAGTGGTGAAGA 1338
QY 241 CCAGTTTTCATACCTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1339 CCAGTTTTCATACCTCATGAGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1398
QY 301 TAATATTCTCAATTTGGGAAGTAGCTGATGTTGGAACAGGAAATTTATCAGAGATGAAGA 360
Db 1399 TAATATTCTCAATTTGGGAAGTAGCTGATGTTGGAACAGGAAATTTATCAGAGATGAAGA 1458
QY 361 AACTGGAAGTACAAGCAGATGAATCTCTTAATTAAGATGGAGTGGCTCAGGGTAGC 420
Db 1459 AACTGGAAGTACAAGCAGATGAATCTCTTAATTAAGATGGAGTGGCTCAGGGTAGC 1518
QY 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTT----- 454
Db 1519 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTT----- 1578
QY 455 ----- 454

Db 1579 TACTCAGACTGTACTCTGTGTGACACAACCTGTGGTTACTAAGGAAAACCTGCCATCTCCAA 1638
QY 455 -----TTAATGGATCTCCAGAAATCAGAA 477
Db 1639 ACTAGAAATGCCATCTTCCCTTGATGTTGGAGCATAGATTTACTGCACAGTTCCCCCTGGA 1698
QY 478 ACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAAAGAAACAAAGGAAAATGGAGGA 537
Db 1699 CCTGGAAGAGTTTCTTGCTGGCTTACAGAAAGCTGAAACCAACTGCTCAATGTCTTACAGGA 1758
QY 538 AGAGCTCTTGGACCTGATCTTGAAGACCTTAAACGCGCAAGTACAAACATATAGAGTGCT 597
Db 1759 TGCTACCGGTAAAGAAAGGCTCTTAGAAGACTCCAGGGAGTAAAGAGCTGTATGAACA 1818
QY 598 TCAAGAAAGATCTAGAACAAAGAAAGTCAAGGCTCAATTTCTCTCATCTCAGATGGTGGT 657
Db 1819 ATGGCAAGACCTCCAAGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGA 1878
QY 658 AGTTGATGA-----ATCTAGTGGAGATCAGCAACTGCTCTTGGAGACA 705
Db 1879 AAACAGCCAAAATACTCTGAGATCCCTGGAGGTTCCGATGATCAGTCTCTGTTACAAAG 1938
QY 706 ACTTAAAGG---TATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAAACCGCTG 762
Db 1939 ACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGGGAAAAGTCTCTCAACATTAG 1998
QY 763 GGTCTCTTTTCAAGACAGATTTCTGACAGTGGAGCGTCTGACCTTTCTCTGAGGAACT 822
Db 1999 GTCCCAATTTGGAAGCCAGTTCTGACCAAGTGAAGCGTCTGACCTTTCTCTGAGGAACT 2058
QY 823 TCTGTTGGCTTACAGCTGAAAGATGATGAATTAAGCGCGCAGGCACTTATTGGAGCGGA 882
Db 2059 TCTGTTGGCTTACAGCTGAAAGATGATGAATTAAGCGCGCAGGCACTTATTGGAGCGGA 2118
QY 883 CTTTCCAGCACTCAGAAAGCAGAACGATGATGATGAGGCGCTTCAAGAGGGAATTTGAAAAC 942
Db 2119 CTTTCCAGCACTCAGAAAGCAGAACGATGATGATGAGGCGCTTCAAGAGGGAATTTGAAAAC 2178
QY 943 TAAAGACCTGTATCATGATGATCTTGGAGCTGTAGATATTTCTGACAGAGGAGCC 1002
Db 2179 TAAAGACCTGTATCATGATGATCTTGGAGCTGTAGATATTTCTGACAGAGGAGCC 2238
QY 1003 TTTTGAAGGAGCTAGAGAACTTACACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGC 1062
Db 2239 TTTTGAAGGAGCTAGAGAACTTACACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGC 2298
QY 1063 CCAGATGTCTACTCGGCTTCTACGAAAGAGCGCTGAGAGGTCATTAATGATGAGTGGGAAA 1122
Db 2299 CCAGATGTCTACTCGGCTTCTACGAAAGAGCGCTGAGAGGTCATTAATGATGAGTGGGAAA 2358
QY 1123 ATTGAACCTGCACTCCGCTGACTGCGCAGAGAAAATAGATGAGACCTTGAAGACTCCA 1182
Db 2359 ATTGAACCTGCACTCCGCTGACTGCGCAGAGAAAATAGATGAGACCTTGAAGACTCCA 2418
QY 1183 GGAACTTCAAGAGCCACCGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGTGGATCAA 1242
Db 2419 GGAACTTCAAGAGCCACCGATGAGTGGACCTCAAGCTGCGCCCAAGCTGAGTGGATCAA 2478
QY 1243 GGGATCTTGGCAGCGCTGGGCGATCTCTCATTTGATCTCTTCCAGATCACCTTCGAGAA 1302
Db 2479 GGGATCTTGGCAGCGCTGGGCGATCTCTCATTTGATCTCTTCCAGATCACCTTCGAGAA 2538
QY 1303 AGTCAAGGCACTTCAGAGGAGAAATTTGCGCTCTGAAAGAAAGTGAAGCCAGCTCAATGA 1362
Db 2539 AGTCAAGGCACTTCAGAGGAGAAATTTGCGCTCTGAAAGAAAGTGAAGCCAGCTCAATGA 2598
QY 1363 CTTGCTCGCAGCTTACCCTTTGGGCACTTGGGCACTTCCCTCATTTGATCTCTTCCAGATCACCTTCGAGAA 1422
Db 2599 CTTGCTCGCAGCTTACCCTTTGGGCACTTGGGCACTTCCCTCATTTGATCTCTTCCAGATCACCTTCGAGAA 2658
QY 1423 GGAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTTGGCGCTGAGGAGCCGAGTCAAGCA 1482

Db 2659 GGAAGACCTGAACACCAAGTGAAGCTTCTGAGGTGGCCGTCGAGGACCGAGTCAGCA 2718
Qy 1483 GTGCAATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGACATTTCTTCCAGCTGTGT 1542
Db 2719 GTGCAATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGACATTTCTTCCAGCTGTGT 2778
Qy 1543 CCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGA 1602
Db 2779 CCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCCTACTATATCAACACGA 2838
Qy 1603 GACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGTCTACAGTCTTTAGTGA 1662
Db 2839 GACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGTCTACAGTCTTTAGTGA 2898
Qy 1663 CCTGAATATGTCAGATTTCTGAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGAA 1722
Db 2899 CCTGAATATGTCAGATTTCTGAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGAA 2958
Qy 1723 GGCCCTTTGCTTGGATCTCTTGAGCCTGTGACCTGCATGTGATGCTTTGGACCAACAA 1782
Db 2959 GGCCCTTTGCTTGGATCTCTTGAGCCTGTGACCTGCATGTGATGCTTTGGACCAACAA 3018
Qy 1783 CCTCAAGCAAAATGACAGCCCATGATATCTGAGATTTATTAATTTGACCACTAT 1842
Db 3019 CCTCAAGCAAAATGACAGCCCATGATATCTGAGATTTATTAATTTGACCACTAT 3078
Qy 1843 TTATGACCCCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCTGGATAT 1902
Db 3079 TTATGACCCCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCTGGATAT 3138
Qy 1903 GTGCTGAATGCTGCTGCTGAATTTATGATACGGGACGAAACAGGAGATCCGCTGCT 1962
Db 3139 GTGCTGAATGCTGCTGCTGAATTTATGATACGGGACGAAACAGGAGATCCGCTGCT 3198
Qy 1963 GTCTTTTAAACCTGGATCATTTCCCTGTGTAAGCACATTTGGAGACAAATACAGATA 2022
Db 3199 GTCTTTTAAACCTGGATCATTTCCCTGTGTAAGCACATTTGGAGACAAATACAGATA 3258
Qy 2023 CCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGTGACCGCAGCGCTGGGCTCCT 2082
Db 3259 CCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTGTGACCGCAGCGCTGGGCTCCT 3318
Qy 2083 TCTGCAATGATCTATCCAA 2101
Db 3319 TCTGCAATGATCTATCCAA 3337

RESULT 3
AX538619
LOCUS AX538619 5417 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 39 from Patent WO0229056.
ACCESSION AX538619
VERSION AX538619.1 GI:25271163
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.O.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 39 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
1. .5417
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"
BASE COUNT 1700 a 1192 c 1182 g 1343 t
ORIGIN
Query Match 76.7%; Score 1611.6; DB 6; Length 5417;
Best Local Similarity 85.3%; Pred. No. 0;

Matches 1872; Conservative 0; Mismatches 229; Indels 93; Gaps 3;
Qy 1 GAGCTATGCTACACACAGGCTGCTTATGTCAACACCTCTGACCTCTACACGAGCCCAT 60
Db 1099 GAGCTATGCTACACACAGGCTGCTTATGTCAACACCTCTGACCTCTACACGAGCCCAT 1158
Qy 61 TCCTTCCACAGCATTTTGGAAAGCTCTCTGAAGACAAGTCATTGTCAGTTCATTGATGAGAG 120
Db 1159 TCCTTCCACAGCATTTTGGAAAGCTCTCTGAAGACAAGTCATTGTCAGTTCATTGATGAGAG 1218
Qy 121 TGAAGTAAACCTGGACCGCTTATCAACACGCTTTAGAGAGATTTATCGTGGCTCTTTTC 180
Db 1219 TGAAGTAAACCTGGACCGCTTATCAACACGCTTTAGAGAGATTTATCGTGGCTCTTTTC 1278
Qy 181 TGCCTGAGGACACATTCGAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1279 TGCCTGAGGACACATTCGAAGCAAGGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1338
Qy 241 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCGGGTTGG 300
Db 1339 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCGGGTTGG 1398
Qy 301 TAATATTTCTACAAATTCGGAAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAAGTGAAGA 360
Db 1399 TAATATTTCTACAAATTCGGAAAGTAAGCTGATTGGAACAGGAAAATTTATCAGAAAGTGAAGA 1458
Qy 361 AACTGAAGTACAAAGACAGATGAATCTCTAAATCAAGATGGGAATGCTCAGGGTAGC 420
Db 1459 AACTGAAGTACAAAGACAGATGAATCTCTAAATCAAGATGGGAATGCTCAGGGTAGC 1518
Qy 421 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAAAT 480
Db 1519 TAGCATGGAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAATCAGAAAAT 1578
Qy 481 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAAGAAAAATGGAGAAAGA 540
Db 1579 GAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGAACAAAGAAAAATGGAGAAAGA 1638
Qy 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACACACATAAAGTGCTTCA 600
Db 1639 GCCTCTTGGACCTGATCTTGAAGACCTTAAAGGCCAAGTACACACATAAAGTGCTTCA 1698
Qy 601 AGAAGATCTAGAAACAAGAACAGTCAAGGTCAATTTCTCACTCACATGCTGCTGCTAGT 660
Db 1699 AGAAGATCTAGAAACAAGAACAGTCAAGGTCAATTTCTCACTCACATGCTGCTGCTAGT 1758
Qy 661 TGATGAATCTAGTGGAGATCACGCAACCTGCTGCTTTGGAAAGAACAACTTAAGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACCTGCTGCTTTGGAAAGAACAACTTAAGTATTGGG 1818
Qy 721 AGATCGATGGCAAAACATCTGTAGATGGACAGACCGCTGGGTTCTTTTACAGACAG 780
Db 1819 AGATCGATGGCAAAACATCTGTAGATGGACAGACCGCTGGGTTCTTTTACAGACAG 1878
Qy 781 TTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGACAGAACTTCTGCTGCTGCTAGCT 840
Db 1879 CCCTTCTCAAAATGCAACGCTTCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTT 1938
Qy 841 GAAAGATGATGAATTAAGCCGCGAGGACCTATTGAGGGGCACTTTCCAGACAGTTCAGAA 900
Db 1939 AAAAGAAAGATGCGATGAACAAAGATTACAACTGGCTTTAAAGATCAAAATGAATGTT 1998
Qy 901 GCAGAACGATGTATCAGAGGCC--TTCAAGAGGGAATTTGAAAACTAAAGAACCTGTAATC 958
Db 1999 ATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2058
Qy 959 ATGAGTACTCTTGAGACTGTACGAATATTT-----CTGACAGACAGCGCTTTGGA 1008
Db 2059 GGGCAAACTCTGATTTCTCACTCAACACAGATCTCTTTCAACACATGAAGATAAGTCAGTGAC 2118
Qy 1009 AGGACTAGAGAAACTCTACAGGAGCCGACAGAGCTGCCTCTCTGAGGAGAGAGACCCAGAA 1068
Db 2119 CCAGAACAGGAAGCATGGCTGTGATTAATTTGCCCCGCTGTGGGATAATTTAGTCCAAAA 2178

QY 1069 TGTCACTGGCTTTCTAGAAAGCAGGCTGAGAGGTCAATAC----- 1110
Db 2179 ACTTGAAGAGTACAGCACAGATTTTACAGGCTGTCAACCACTCAGCCATCACTAAC 2238
QY 1111 ----- 1110
Db 2239 ACAGCAACTGTAAATGAAACAGTAACACGTTGACACAGGAAACAGATCTCGTAA 2298
QY 1111 ---TGAAGTGGGAAAAATTTGAACCTCGACTCGCTGACTGCGCAGAGAAAAATAGATGAGAC 1167
Db 2299 GCATGCTCAAGAGAACTTCCACACCACTCCCAAAAGAGAGGAGAGATTAATCTGGA 2358
QY 1168 CTTTGAAGACTCCAGAACTTCAAGAGGCCACGATGAGCTGGAACCTCAAGCTGGGCA 1227
Db 2359 TCTTGAAGACTCCAGAACTTCAAGAGGCCACGATGAGCTGGAACCTCAAGCTGGGCA 2418
QY 1228 AGCTGAGGTGATCAAGGATCTTGGGAGCCGCTGGGAGTCTCTCATTTGACTCTTCCA 1287
Db 2419 AGCTGAGGTGATCAAGGATCTTGGGAGCCGCTGGGAGTCTCTCATTTGACTCTTCCA 2478
QY 1288 AGATCACTCGAGAACTCAAGGACTTCAAGGAGAAATTCGCGCTTCAAGAGAACT 1347
Db 2479 AGATCACTCGAGAACTCAAGGACTTCAAGGAGAAATTCGCGCTTCAAGAGAACT 2538
QY 1348 GAGCAGCTCAATGACCTTCTGCGCAGCTTACACTTTGGGATTCAGCTCTCACCGTA 1407
Db 2539 GAGCAGCTCAATGACCTTCTGCGCAGCTTACACTTTGGGATTCAGCTCTCACCGTA 2598
QY 1408 TAACCTCAGCACTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTGCGCTCGA 1467
Db 2599 TAACCTCAGCACTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTGCGCTCGA 2658
QY 1468 GGACGAGTCAAGGAGCTGATGAGAGCCCAAGGACTTTGGTCCAGATCTCAGCACTT 1527
Db 2659 GGACGAGTCAAGGAGCTGATGAGAGCCCAAGGACTTTGGTCCAGATCTCAGCACTT 2718
QY 1528 TCTTTCCAGCTCTGTCAGGCTTCTGCGGAGAGGCTTCTGCGGAGGCTTCTGCGGAGGCTT 1587
Db 2719 TCTTTCCAGCTCTGTCAGGCTTCTGCGGAGAGGCTTCTGCGGAGGCTTCTGCGGAGGCTT 2778
QY 1588 CTATATCAACCAAGAGCTCAAACTTTGCTGGGACCATCCCAAAATGACAGAGCTCTA 1647
Db 2779 CTATATCAACCAAGAGCTCAAACTTTGCTGGGACCATCCCAAAATGACAGAGCTCTA 2838
QY 1648 CCAGTCTTAGCTGACCTGAATATGTCAATTTCTAGCTTATAGGACTGCCATGAACT 1707
Db 2839 CCAGTCTTAGCTGACCTGAATATGTCAATTTCTAGCTTATAGGACTGCCATGAACT 2898
QY 1708 CCGAAGACTGCAAGAGGCTTTGCTTTGGATCTCTTGAAGCTTCTGAGCTGCTGATGC 1767
Db 2899 CCGAAGACTGCAAGAGGCTTTGCTTTGGATCTCTTGAAGCTTCTGAGCTGCTGATGC 2958
QY 1768 CTTGGACGAGCAAACTCAAGCAAAATGACAGCCCATGATATCTGAGATTTAA 1827
Db 2959 CTTGGACGAGCAAACTCAAGCAAAATGACAGCCCATGATATCTGAGATTTAA 3018
QY 1828 TTGTTTGAACCACTTTATGACCCCTTGGAGCAAGCAAACTTTGCTCAACCTCC 1887
Db 3019 TTGTTTGAACCACTTTATGACCCCTTGGAGCAAGCAAACTTTGCTCAACCTCC 3078
QY 1888 TCTCTGGGTGATATGTCTGAATCTGGCTGCTGAATGTTTATGATACGGGAGCAAG 1947
Db 3079 TCTCTGGGTGATATGTCTGAATCTGGCTGCTGAATGTTTATGATACGGGAGCAAG 3138
QY 1948 GAGGATCGGTCTGCTTTTAACTGGATCATTTCCCTGTTGTAAGCAATTTGA 2007
Db 3139 GAGGATCGGTCTGCTTTTAACTGGATCATTTCCCTGTTGTAAGCAATTTGA 3198
QY 2008 AGACAAGTACAGATCTTTTCAAGCAAGTGGCAAGTTTCAAGAGATTTTGTGACCAAG 2067
Db 3199 AGACAAGTACAGATCTTTTCAAGCAAGTGGCAAGTTTCAAGAGATTTTGTGACCAAG 3258

QY 2068 CAGCTGGGCTCTCTTCTGATGATTTCTATCCAA 2101
Db 3259 CAGCTGGGCTCTCTTCTGATGATTTCTATCCAA 3292
RESULT 4
AR304538 AR304538 5952 bp DNA linear PAT 12-JUN-2003
LOCUS Sequence 1 from patent US 6544786.
DEFINITION AR304538
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
source /organism="unknown"
1. 5952
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 63.2%; Score 1328.8; DB 6; Length 5952;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 762 GGGTCTTTTACAGACAGATTTCTGACGAGTGGAGCGTCTGACCTTTCTCTGAGAAC 821
Db 3284 GGTCCCAATTTGGAAGCCAGTTCTGACGAGTGGAGCGTCTGACCTTTCTCTGAGAAC 3343
QY 822 TTCGGTGTGCTACAGCTGAAAGATGATGATTAATTAAGCCGCGACGACCTATTGAGGCG 881
Db 3344 TTCGGTGTGCTACAGCTGAAAGATGATGATTAATTAAGCCGCGACGACCTATTGAGGCG 3403
QY 882 ACTTTCACGAGCTTCAAGAGCAGAACCATGTACATAGGGCTTCAAGAGGGAATTTGAAA 941
Db 3404 ACTTTCACGAGCTTCAAGAGCAGAACCATGTACATAGGGCTTCAAGAGGGAATTTGAAA 3463
QY 942 CTAAGAACCTTGTATCATAGTACTTCTGAGACTGTACGATATTTCTGACAGAGCAGC 1001
Db 3464 CTAAGAACCTTGTATCATAGTACTTCTGAGACTGTACGATATTTCTGACAGAGCAGC 3523
QY 1002 CTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGAGAGAG 1061
Db 3524 CTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGAGAGAGAG 3583
QY 1052 CCCAGAACTGCTGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAA 1121
Db 3584 CCCAGAACTGCTGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAA 3643
QY 1122 AATTGAACCTGCTCTGCTGACTGGCAGAGAAATAGATGAGACCTTCAAGAGACTCC 1181
Db 3644 AATTGAACCTGCTCTGCTGACTGGCAGAGAAATAGATGAGACCTTCAAGAGACTCC 3703
QY 1182 AGGAACCTTCAAGAGCCACCGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTATCA 1241
Db 3704 AGGAACCTTCAAGAGCCACCGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTATCA 3763
QY 1242 AGGGATCTCTGAGCCCTGGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGA 1301
Db 3764 AGGGATCTCTGAGCCCTGGGCGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGA 3823
QY 1302 AAGTCAAGGCACTTCCAGAGAGAAATTTGGCTCTCTGAAAGAGAACTGAGCCCTCAATG 1361
Db 3824 AAGTCAAGGCACTTCCAGAGAGAAATTTGGCTCTCTGAAAGAGAACTGAGCCCTCAATG 3883
QY 1362 ACCTTGTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCAACGATTAACCTCAGCACTC 1421
Db 3884 ACCTTGTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCAACGATTAACCTCAGCACTC 3943

QY 1422 TGGAGAGCCTGAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGC 1481
DB 3944 TGGAGAGCCTGAACACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGC 4003
QY 1482 AGCTGCATGAAGCCACACAGGACATTTGCTCCAGCATCTCAGCATCTTCTTCCACAGTCG 1541
DB 4004 AGCTGCATGAAGCCACACAGGACATTTGCTCCAGCATCTCAGCATCTTCTTCCACAGTCG 4063
QY 1542 TCCAGGTCCTTGGAGAGAGCCATCTCGCCAAAGTGGCCCTACTATATCAACACG 1601
DB 4064 TCCAGGTCCTTGGAGAGAGCCATCTCGCCAAAGTGGCCCTACTATATCAACACG 4123
QY 1602 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGCTG 1661
DB 4124 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGCTCTTTAGCTG 4183
QY 1662 ACCTGAATAATGTCTAGATCTCAGCTTATAGCATGCGCATGAACCTCCGAAGACTGCAGA 1721
DB 4184 ACCTGAATAATGTCTAGATCTCAGCTTATAGCATGCGCATGAACCTCCGAAGACTGCAGA 4243
QY 1722 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCGATGATGCTTGGACAGCACA 1781
DB 4244 AGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCGATGATGCTTGGACAGCACA 4303
QY 1782 ACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTTAATTTGTTGACCACTA 1841
DB 4304 ACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTTAATTTGTTGACCACTA 4363
QY 1842 TTATGACCGCTGAGCAAGACACACAAATTTGTCACAGCTCCCTCTGCTGGATGATA 1901
DB 4364 TTATGACCGCTGAGCAAGACACACAAATTTGTCACAGCTCCCTCTGCTGGATGATA 4423
QY 1902 TGTGTCTGAACCTGGCTGTGTAATGTTATGATACGGGACGACAGGAGGATCGGTGTC 1961
DB 4424 TGTGTCTGAACCTGGCTGTGTAATGTTATGATACGGGACGACAGGAGGATCGGTGTC 4483
QY 1962 TGTCTTTTAAACTGGCATCTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT 2021
DB 4484 TGTCTTTTAAACTGGCATCTTCCCTGTGTAAAGCACATTTGGAAGACAAAGTACAGAT 4543
QY 2022 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACGAGCGAGGCTGGGCTCC 2081
DB 4544 ACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACGAGCGAGGCTGGGCTCC 4603
QY 2082 TTCTGCATGATCTATCCAA 2101
DB 4604 TTCTGCATGATCTATCCAA 4623

RESULT 5
AX114289
LOCUS 5952 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent WO0129243.
ACCESSION AX114289
VERSION AX114289.1 GI:14031259
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Paul X.L. and Xiao, X.
TITLE Method and vector for producing and transferring trans -spliced
peptides
JOURNAL Patent: WO 0129243-A 1 26-APR-2001;
DALHOUSIE UNIVERSITY (CA) ; UNIV. OF PITTSBURGH OF THE COMMONWEALTH
SYSTEM OF HIGHER EDUCATION (US)
FEATURES
Location/Qualifiers
1..5952
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

misc_feature 2897..2898
/note="S4 junction site"
misc_feature 3198..3199
/note="S2 junction site"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN
Query Match 63.2%; Score 1328.8; DB 6; Length 5952;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 762 GGGTCTTTTACAGACAGTCTGACAGTGGAGCGTCTGACCTTCTCTGAGGAGC 821
DB 3284 GGTCCCATTTGGAAGCCAGTTCTGACCACTGGAAGCGTCTGACCTTCTCTGAGGAGC 3343
QY 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCAGCCTATTGGAGCG 881
DB 3344 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCAGCCTATTGGAGCG 3403
QY 882 ACTTCCAGCAGTTTCAAGAGCAAGATGATAGAGGCTTCAAGAGGCAATTTGAAA 941
DB 3404 ACTTCCAGCAGTTTCAAGAGCAAGATGATAGAGGCTTCAAGAGGCAATTTGAAA 3463
QY 942 CTAAAGAACCTGTAATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 1001
DB 3464 CTAAAGAACCTGTAATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGC 3523
QY 1002 CTTTGGAGAGCTTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG 1061
DB 3524 CTTTGGAGAGCTTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAG 3583
QY 1062 CCCAGATGCTCACTGGCTTCTACGAAAGCAGGCTGAGAGGTCAATCTGAGTGGGAAA 1121
DB 3584 CCCAGATGCTCACTGGCTTCTACGAAAGCAGGCTGAGAGGTCAATCTGAGTGGGAAA 3643
QY 1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAGACTCC 1181
DB 3644 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAGACTCC 3703
QY 1182 AGGAATCTCAAGAGCCAGGATGAGCTCAAGCTGCGCCCAAGCTGAGGTGATCA 1241
DB 3704 AGGAATCTCAAGAGCCAGGATGAGCTCAAGCTGCGCCCAAGCTGAGGTGATCA 3763
QY 1242 AGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACTTCGAGA 1301
DB 3764 AGGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACTTCGAGA 3823
QY 1302 AAGTCAAGGACCTTCAGAGAGAAATTTGGCCCTCTGAAAGAGAACCTGAGCCAGCTCAATG 1361
DB 3824 AAGTCAAGGACCTTCAGAGAGAAATTTGGCCCTCTGAAAGAGAACCTGAGCCAGCTCAATG 3883
QY 1362 ACCTTCTCGCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATATACCTCAGCAGCTC 1421
DB 3884 ACCTTCTCGCAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATATACCTCAGCAGCTC 3943
QY 1422 TGAAGACCTTGAACACCAAGATGGAAGTTTCTGAGGTGGCCCTCGAGGACCGAGTCAGGC 1481
DB 3944 TGAAGACCTTGAACACCAAGATGGAAGTTTCTGAGGTGGCCCTCGAGGACCGAGTCAGGC 4003
QY 1482 AGTGCATGAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCAGCTTTCTTCCAGCTCTG 1541
DB 4004 AGTGCATGAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCAGCTTTCTTCCAGCTCTG 4063
QY 1542 TCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAAGTGGCCCTACTATATCAACACG 1601
DB 4064 TCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAAGTGGCCCTACTATATCAACACG 4123
QY 1602 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACAGCTCTTTAGCTG 1661
DB 4124 AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACAGCTCTTTAGCTG 4183
QY 1662 ACCTGAATAATGTCAAGTTCTCAGCTTATAGGAGCTGCCATGAACTCCGAAGACTGCAGA 1721

Db	4184	ACCTGAAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAATCCGAGACTGCAGA	4249
Qy	1722	AGGCCCTTTGCTTTGGATCTCTTGAGCCCTGTTCAGCTGCATGTGATGCTTTCGACCAGCAC	1781
Db	4244	AGGCCCTTTGCTTTGGATCTCTTGAGCCCTGTTCAGCTGCATGTGATGCTTTCGACCAGCAC	4303
Qy	1782	ACCTCAAGCAAAATGACAGCCCATGATATCTTGACAGATATTAATTTGTTGACCACTA	1841
Db	4304	ACCTCAAGCAAAATGACAGCCCATGATATCTTGACAGATATTAATTTGTTGACCACTA	4363
Qy	1842	TTTATGACCCCTTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGATA	1901
Db	4364	TTTATGACCCCTTGGAGCAAGAGCACAAATTTGGTCAACGTCCTCTCTGCGTGGATA	4423
Qy	1902	TGTGTCGAATCGCTCTCTGAAATTTATGATACGGGACGAAACAGGAGGATCCGTGTCC	1961
Db	4424	TGTGTCGAATCGCTCTCTGAAATTTATGATACGGGACGAAACAGGAGGATCCGTGTCC	4483
Qy	1962	TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTTGGAAAGCAAGTACAGAT	2021
Db	4484	TGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTTGGAAAGCAAGTACAGAT	4543
Qy	2022	ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTCGGCTCC	2081
Db	4544	ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTCGGCTCC	4603
Qy	2082	TTCTGCATGATTTCTATCCAA	2101
Db	4604	TTCTGCATGATTTCTATCCAA	4623

RESULT 6

AX538622

LOCUS AX538622 Sequence 42 from Patent WO0229056.

DEFINITION AX538622

ACCESSION AX538622

VERSION AX538622.1 GI:25271171

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Chamberlain,J.S. and Harper,S.Q.

TITLE Mini-dystrophin nucleic acid and peptide sequences

JOURNAL Patent: WO 0229056-A 42 11-APR-2002;

THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES

source Location/Qualifiers

1..8689

BASE COUNT 2721 a 1804 c 1861 g 2303 t

ORIGIN

Query Match 63.2%; Score 1328.8; DB 6; Length 8689;

Best Local Similarity 99.5%; Pred.No. 0;

Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 762 GGGTTCTTTTACAGACAGTCTTGACCAAGTGGAAAGCGTCTGCACCTTCTCTCGAGGAAC 821

Db 3330 GGTCCCAATTTGGAAGCCAGTCTTGACCAAGTGGAAAGCGTCTGCACCTTCTCTCGAGGAAC 3389

Qy 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAAATTAAGCCGGCAGGCACTATTGGAGCG 881

Db 3390 TTCTGGTGTGGCTACAGCTGAAAGATGATGAAATTAAGCCGGCAGGCACTATTGGAGCG 3449

Qy 882 ACTTTCAGCAGTTTCAGAACAGAACCATGTACATAGGGCTTTCAGAGGGAATTTGAAA 941

Db 3450 ACTTTCAGCAGTTTCAGAACAGAACCATGTACATAGGGCTTTCAGAGGGAATTTGAAA 3509

Qy 942 CTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGCACAGACGC 1001

3510	CTAAGAA	CGCTGTAATCA	TGAGTAC	TCTTTGAGACTGTA	CGAATATTTTCTGACAGACGACG	3569
Qy	CTTTTGAAGGACTAGAGAA	ACTCTACAGAGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAG	1061			
	CTTTTGAAGGACTAGAGAA	ACTCTACAGAGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAG	3629			
	CTTTTGAAGGACTAGAGAA	ACTCTACAGAGAGCCAGAGAGCTGCCTCTCTGAGGAGAGAG	3629			
Qy	CCAGAAATGTCAC	TCCGCTTCTACGAAACAGGCTGAGGAGTCA	TACTAGTGGGAA	1121		
	CCAGAAATGTCAC	TCCGCTTCTACGAAACAGGCTGAGGAGTCA	TACTAGTGGGAA	3689		
	CCAGAAATGTCAC	TCCGCTTCTACGAAACAGGCTGAGGAGTCA	TACTAGTGGGAA	3689		
Qy	AATTGAACTGCACT	CCGCTGACTGGCAGAGAAAAATAGATCAGAC	CCCTTGAAAGACTCC	1181		
	AATTGAACTGCACT	CCGCTGACTGGCAGAGAAAAATAGATCAGAC	CCCTTGAAAGACTCC	3749		
	AATTGAACTGCACT	CCGCTGACTGGCAGAGAAAAATAGATCAGAC	CCCTTGAAAGACTCC	3749		
Qy	AGGAACTTCAAGAGGCCA	CGATAGACTGGA	CCTCAAGCTGGCCNAGCTGAGTGTATCA	1241		
	AGGAACTTCAAGAGGCCA	CGATAGACTGGA	CCTCAAGCTGGCCNAGCTGAGTGTATCA	3809		
	AGGAACTTCAAGAGGCCA	CGATAGACTGGA	CCTCAAGCTGGCCNAGCTGAGTGTATCA	3809		
Qy	AGGATCTCTGGCAGCCG	TGGCGATCTCCTCATTGACTCTCTCAAAGATCA	CCCTCGAGA	1301		
	AGGATCTCTGGCAGCCG	TGGCGATCTCCTCATTGACTCTCTCAAAGATCA	CCCTCGAGA	3869		
	AGGATCTCTGGCAGCCG	TGGCGATCTCCTCATTGACTCTCTCAAAGATCA	CCCTCGAGA	3869		
Qy	AGTCAAGGACATT	CGAGAGAAATTCGCTCTGAAAGAGAA	CGTAGCAGTCAATG	1361		
	AGTCAAGGACATT	CGAGAGAAATTCGCTCTGAAAGAGAA	CGTAGCAGTCAATG	3929		
	AGTCAAGGACATT	CGAGAGAAATTCGCTCTGAAAGAGAA	CGTAGCAGTCAATG	3929		
Qy	ACCTTCTGCTGCCAGCTT	ACCATTTTGGGATTCAGCTCTCACCGTATA	ACCTCAGCACATC	1421		
	ACCTTCTGCTGCCAGCTT	ACCATTTTGGGATTCAGCTCTCACCGTATA	ACCTCAGCACATC	3989		
	ACCTTCTGCTGCCAGCTT	ACCATTTTGGGATTCAGCTCTCACCGTATA	ACCTCAGCACATC	3989		
Qy	TGGAAGACCTTGAA	CACAGATGGAAGCTTCTCGAGGTGCCCTCGAGGACCGAGTCAGGC	1481			
	TGGAAGACCTTGAA	CACAGATGGAAGCTTCTCGAGGTGCCCTCGAGGACCGAGTCAGGC	4049			
	TGGAAGACCTTGAA	CACAGATGGAAGCTTCTCGAGGTGCCCTCGAGGACCGAGTCAGGC	4049			
Qy	AGCTGATGAGCCAC	AGGACTTTTGGTCCAGCATCTCAGACATCTTCTTCCAGCTCG	1541			
	AGCTGATGAGCCAC	AGGACTTTTGGTCCAGCATCTCAGACATCTTCTTCCAGCTCG	4109			
	AGCTGATGAGCCAC	AGGACTTTTGGTCCAGCATCTCAGACATCTTCTTCCAGCTCG	4109			
Qy	TCCAGGCTCCCTGGGAGAG	AGCCATCTCGGCCAAACAAAGTGCCCTACTATATCAACCAAG	1601			
	TCCAGGCTCCCTGGGAGAG	AGCCATCTCGGCCAAACAAAGTGCCCTACTATATCAACCAAG	4169			
	TCCAGGCTCCCTGGGAGAG	AGCCATCTCGGCCAAACAAAGTGCCCTACTATATCAACCAAG	4169			
Qy	AGACTCAAA	CAACTTCTGGGACCATCCCAAATGACAGAGCTTACAGTCTTTAGCTG	1661			
	AGACTCAAA	CAACTTCTGGGACCATCCCAAATGACAGAGCTTACAGTCTTTAGCTG	4229			
	AGACTCAAA	CAACTTCTGGGACCATCCCAAATGACAGAGCTTACAGTCTTTAGCTG	4229			
Qy	ACCTGAAATATGTCAGAT	TCTCAGCTTATAGAGACTGCCATGAACTCCGAAAGCTGCAGA	1721			
	ACCTGAAATATGTCAGAT	TCTCAGCTTATAGAGACTGCCATGAACTCCGAAAGCTGCAGA	4289			
	ACCTGAAATATGTCAGAT	TCTCAGCTTATAGAGACTGCCATGAACTCCGAAAGCTGCAGA	4289			
Qy	AGGCCCTTTGCTTGGAT	CTCTTGGAGCCTGTACGCTGCA	TGTGATGCTTGGACCAAGCAC	1781		
	AGGCCCTTTGCTTGGAT	CTCTTGGAGCCTGTACGCTGCA	TGTGATGCTTGGACCAAGCAC	4349		
	AGGCCCTTTGCTTGGAT	CTCTTGGAGCCTGTACGCTGCA	TGTGATGCTTGGACCAAGCAC	4349		
Qy	ACCTCAAGCAAAATGAC	AGCCCATGATCTCTGAGACTTGAATCTTTGACCACTA	1841			
	ACCTCAAGCAAAATGAC	AGCCCATGATCTCTGAGACTTGAATCTTTGACCACTA	4409			
	ACCTCAAGCAAAATGAC	AGCCCATGATCTCTGAGACTTGAATCTTTGACCACTA	4409			
Qy	TTTTATGAC	CCCTTGGAGCAGCAACAAATTTGGTCAACGTCCTCTCTCGTGGATA	1901			
	TTTTATGAC	CCCTTGGAGCAGCAACAAATTTGGTCAACGTCCTCTCTCGTGGATA	4469			
	TTTTATGAC	CCCTTGGAGCAGCAACAAATTTGGTCAACGTCCTCTCTCGTGGATA	4469			
Qy	TGTGTCTGA	ACTGGCTGTGAATGTTATACGGGACGAAACAGGGAGGATCCCGTCTCC	1961			
	TGTGTCTGA	ACTGGCTGTGAATGTTATACGGGACGAAACAGGGAGGATCCCGTCTCC	4529			
	TGTGTCTGA	ACTGGCTGTGAATGTTATACGGGACGAAACAGGGAGGATCCCGTCTCC	4529			
Qy	TGTCTTTTAAA	CTGGCATCAITTCCTGTGTAAGCAATTTGGAAGCAAGTACAGAT	2021			
	TGTCTTTTAAA	CTGGCATCAITTCCTGTGTAAGCAATTTGGAAGCAAGTACAGAT	4589			
	TGTCTTTTAAA	CTGGCATCAITTCCTGTGTAAGCAATTTGGAAGCAAGTACAGAT	4589			
Qy	ACCTTTTCAAGCAAGTGG	CAAGTTCAACAGGATTTTGTGACCAAGCGCAGCTGGGCTCC	2081			
	ACCTTTTCAAGCAAGTGG	CAAGTTCAACAGGATTTTGTGACCAAGCGCAGCTGGGCTCC	4649			
	ACCTTTTCAAGCAAGTGG	CAAGTTCAACAGGATTTTGTGACCAAGCGCAGCTGGGCTCC	4649			

2082 TTCTGCATGATTCTATCCAA 2101
4650 TTCTGCATGATTCTATCCAA 4669

RESULT 7
AX538624
LOCUS
DEFINITION
Sequence 44 from Patent WO0229056.
ACCESSION
AX538624
VERSION
AX538624.1 GI:25271175
KEYWORDS
synthetic construct
SOURCE
synthetic construct
ORGANISM
artificial sequences.

REFERENCE
1
AUTHORS
Chamberlain, J.S. and Harper, S.O.
TITLE
Mini-dystrophin nucleic acid and peptide sequences
JOURNAL
Patent: WO 0229056-A 44 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..1143
source
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"

BASE COUNT 3707 a 2339 c 2502 g 2895 t

ORIGIN
Query Match 63.2%; Score 1328.8; DB 6; Length 11443;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

762 GGCTTTCTTTACAGACAGTTCTGACAGTGGAGGCTCTGCACCTTTCTCTGCGAGGAC 821
6084 GTGCCAATTTGGAGCGAGTTCTGACAGTGGAGGCTCTGCACCTTTCTCTGCGAGGAC 6143
822 TTCTGGTGTGGCTACAGTGAAGATGATGAATTAAGCCGCGAGGACCTATTGAGGCG 881
6144 TTCTGGTGTGGCTACAGTGAAGATGATGAATTAAGCCGCGAGGACCTATTGAGGCG 6203
882 ACTTCCAGCAGTTCAGAACGACGATGATAGAGGCTTCAGAGGGGAATGAAGA 941
6204 ACTTCCAGCAGTTCAGAACGACGATGATAGAGGCTTCAGAGGGGAATGAAGA 6263
942 CTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTAGGAATATTCTGACAGAGCAGC 1001
6264 CTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTAGGAATATTCTGACAGAGCAGC 6323
1002 CTTTGGAGGACTAGAGAACTCTACAGAGCCGAGAGCTGCTCTCTGAGGAGAGAG 1061
6324 CTTTGGAGGACTAGAGAACTCTACAGAGCCGAGAGCTGCTCTCTGAGGAGAGAG 6383
1062 CCAGAAATGTCACCTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 1121
6384 CCAGAAATGTCACCTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAA 6443
1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAATGATGAGACCCCTTGAAGACTCC 1181
6444 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAATAATGATGAGACCCCTTGAAGACTCC 6503
1182 AGGAACCTCAAGAGGCCAGCGATGAGTGGACCTCAAGCTGCGCCAGCTGAGGTGATCA 1241
6504 AGGAACCTCAAGAGGCCAGCGATGAGTGGACCTCAAGCTGCGCCAGCTGAGGTGATCA 6563
1242 AGGGATCTGCGACCGCTGGGCGATCTCTCAATGACTCTCTCCAGATCACCTCGAGA 1301
6564 AGGGATCTGCGACCGCTGGGCGATCTCTCAATGACTCTCTCCAGATCACCTCGAGA 6623
1302 AAGTCAAGGCACTTCGAGGAGAAATTTGGGCTCTGAAGAGACGTCGAGCCAGTCAATG 1361
6624 AAGTCAAGGCACTTCGAGGAGAAATTTGGGCTCTGAAGAGACGTCGAGCCAGTCAATG 6683

1362 ACCTTGCTCGCAGCTTACCACTTTGGGCATTACAGCTCTCACCGTATACCTCAGCACTC 1421
6684 ACCTTGCTCGCAGCTTACCACTTTGGGCATTACAGCTCTCACCGTATACCTCAGCACTC 6743
1422 TGAAGACCTGAACACACAGATGGAAGCTTCTCGAGTGCCCTCGAGGACCCAGTCCAGC 1481
6744 TGAAGACCTGAACACACAGATGGAAGCTTCTCGAGTGCCCTCGAGGACCCAGTCCAGC 6803
1482 AGCTGCATGAAGCCACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTCG 1541
6804 AGCTGCATGAAGCCACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTCG 6863
1542 TCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 1601
6864 TCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 6923
1602 AGACTCAACAACTTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 1661
6924 AGACTCAACAACTTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 6983
1662 ACCTGAATATGTCAGATTTCTCAGCTTTATAGGACTGCCATGAACTCCGAAAGACTCCAGA 1721
6984 ACCTGAATATGTCAGATTTCTCAGCTTTATAGGACTGCCATGAACTCCGAAAGACTCCAGA 7043
1722 AGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGATGATGCTTTGGACCACTA 1781
7044 AGCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGATGATGCTTTGGACCACTA 7103
1782 ACCTCAAGCAAAATGACAGCCCATGATATCTGCAGATTTATTAATTTGTTGACCACTA 1841
7104 ACCTCAAGCAAAATGACAGCCCATGATATCTGCAGATTTATTAATTTGTTGACCACTA 7163
1842 TTTATGACCCCTGGAGCAAGACACAAATTTGTCGTCGTCCTCTCTGCGTGGATA 1901
7164 TTTATGACCCCTGGAGCAAGACACAAATTTGTCGTCGTCCTCTCTGCGTGGATA 7223
1902 TGTGCTGAACTGGCTGCTGAATGTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 1961
7224 TGTGCTGAACTGGCTGCTGAATGTTATGATACGGGACGAAACAGGAGGATCCGTGTCC 7283
1962 TGTCTTTTAAACCTGGCATCATTTCCCTGTTGTAAGACACATTTGGAAGACAACTACAGAT 2021
7284 TGTCTTTTAAACCTGGCATCATTTCCCTGTTGTAAGACACATTTGGAAGACAACTACAGAT 7343
2022 ACCTTTCAAGCAAGTGGAGTCAAGTTCACAGGATTTGTGACCAAGCGCAGCTCGGCTCC 2081
7344 ACCTTTTCAAGCAAGTGGAGTTCACAGGATTTGTGACCAAGCGCAGCTCGGCTCC 7403
2082 TTCTGCATGATTCTATCCAA 2101
7404 TTCTGCATGATTCTATCCAA 7423

RESULT 8
AX538627
LOCUS
DEFINITION
Sequence 47 from Patent WO0229056.
ACCESSION
AX538627
VERSION
AX538627.1 GI:25271181
KEYWORDS
synthetic construct
SOURCE
synthetic construct
ORGANISM
artificial sequences.

REFERENCE
1
AUTHORS
Chamberlain, J.S. and Harper, S.O.
TITLE
Mini-dystrophin nucleic acid and peptide sequences
JOURNAL
Patent: WO 0229056-A 47 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..12057
source
/organism="synthetic construct"
/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="Synthetic"

BASE COUNT 4020 a 2456 c 2790 g 2791 t

ORIGIN

Query Match 63.2%; Score 1328.8; DB 6; Length 12057;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
2Y 762 GGGTCTCTTTTACAGACAGTCTGACAGTGAAGCGTCTGCACCTTCTCTGCAGGAAC 821
Db 8598 GGTCCCATTTGGAGCGAGTTCTGACAGTGNAGGCTCTGCACCTTCTCTGCAGGAAC 8657
2Y 822 TCTGTGTGGCTTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACTATTTGGAGCG 881
Db 8658 TCTGTGTGGCTTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACTATTTGGAGCG 8717
2Y 882 ACTTCCAGAGTTCAGAGCAGAGATGTACATAGGCGCTTCAAGAGGGAATTTGAAA 941
Db 8718 ACTTCCAGAGTTCAGAGCAGAGATGTACATAGGCGCTTCAAGAGGGAATTTGAAA 8777
2Y 942 CTAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACGAC 1001
Db 8778 CTAAGAACTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACGAC 8837
2Y 1002 CTTTGGAGAGCTAGAGAACTCTACAGAGGCCAGAGAGTGCCTCTCTGAGGAGAG 1061
Db 8838 CTTTGGAGAGCTAGAGAACTCTACAGAGGCCAGAGAGTGCCTCTCTGAGGAGAGAG 8897
2Y 1062 CCAGAGATGTCACTCGCGCTTCTACGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAA 1121
Db 8898 CCAGAGATGTCACTCGCGCTTCTACGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAA 8957
2Y 1122 AATTGAACCTGCACTCGCTGACTGCGCAGAGAAAATAGATGAGACCTTTGAAAGTCC 1181
Db 8958 AATTGAACCTGCACTCGCTGACTGCGCAGAGAAAATAGATGAGACCTTTGAAAGTCC 9017
2Y 1182 AGGAATCTCAAGAGGCCAGGATGAGTCAAGCTCAAGCTGCGCAAGCTGAGGTGATCA 1241
Db 9018 AGGAATCTCAAGAGGCCAGGATGAGTCAAGCTCAAGCTGCGCAAGCTGAGGTGATCA 9077
2Y 1242 AGGATCTCTGGCAGCCGCTGGGCGATCTCTCTATTGACTCTCTTCCAAGATCACTCGAGA 1301
Db 9078 AGGATCTCTGGCAGCCGCTGGGCGATCTCTCTATTGACTCTCTTCCAAGATCACTCGAGA 9137
2Y 1302 AGTCAAGGACCTCGAGGAGAAATTCGGCTCTGAAAGAGACGTCGAGCCAGTCAATG 1361
Db 9138 AAGTCAAGGACCTCGAGGAGAAATTCGGCTCTGAAAGAGACGTCGAGCCAGTCAATG 9197
2Y 1362 ACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
Db 9198 ACCTTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAACCTCAGCACTC 9257
2Y 1422 TGAAGACCTTGAACACAGATGGAAGTTCTGAGGTGGCGCTGAGGACCGAGTCAGGC 1481
Db 9258 TGAAGACCTTGAACACAGATGGAAGTTCTGAGGTGGCGCTGAGGACCGAGTCAGGC 9317
2Y 1482 AGTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTG 1541
Db 9318 AGTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTG 9377
2Y 1542 TCCAGGCTCCCTGGGAGAGACCATTCGCCAAACAAAGTGCCTACTATATCAACACG 1601
Db 9378 TCCAGGCTCCCTGGGAGAGACCATTCGCCAAACAAAGTGCCTACTATATCAACACG 9437
2Y 1602 AGACTCAACAACTTGTCTGGGACCATCCCAATGACAGAGCTCTACGAGCTTTAGCTG 1661
Db 9438 AGACTCAACAACTTGTCTGGGACCATCCCAATGACAGAGCTCTACGAGCTTTAGCTG 9497
2Y 1662 ACCTGGAATATGTGAGTATTCAGCTTATAGGAGTGCATGAAACTCCGAGAGCTGCAGA 1721
Db 9498 ACCTGGAATATGTGAGTATTCAGCTTATAGGAGTGCATGAAACTCCGAGAGCTGCAGA 9557
```

```
QY 1722 AGGCCCTTTGCTTGGATCTCTTGGAGCTGTCACTGATGTGATGCTTGGACAGCACA 1781
Db 9558 AGGCCCTTTGCTTGGATCTCTTGGAGCTGTCACTGATGTGATGCTTGGACAGCACA 9617
QY 1782 ACCTCAAGCAAAATGACAGCCCATGGATATCTTGCAGATTATTAATTGTTGACCACTA 1841
Db 9618 ACCTCAAGCAAAATGACAGCCCATGGATATCTTGCAGATTATTAATTGTTGACCACTA 9677
QY 1842 TTTATGACCCCTGGAGCAAGACCAACAATTGGTCAACGCTCCCTCTCTGCGTGGATA 1901
Db 9678 TTTATGACCCCTGGAGCAAGACCAACAATTGGTCAACGCTCCCTCTCTGCGTGGATA 9737
QY 1902 TGTGTCTGAAGTGGCTGCTGAATGTTTATGATACGGAGCAACAGGAGGATCGTCTCC 1961
Db 9738 TGTGTCTGAAGTGGCTGCTGAATGTTTATGATACGGAGCAACAGGAGGATCGTCTCC 9797
QY 1962 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAGAT 2021
Db 9798 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAGAT 9857
QY 2022 ACCTTTTCAAGCAAGTGGAGTTCACAGGATTTGTGACAGCGCAGGCTGGGCTCC 2081
Db 9858 ACCTTTTCAAGCAAGTGGAGTTCACAGGATTTGTGACAGCGCAGGCTGGGCTCC 9917
QY 2082 TTTCTGATGATTCTATCCAA 2101
Db 9918 TTTCTGATGATTCTATCCAA 9937
```

RESULT 9

```
HSDMDR Human mRNA for dystrophin. 12446 bp mRNA linear PRI 12-SEP-1993
LOCUS X14298
ACCESSION X14298
VERSION X14298.1 GI:30845
KEYWORDS Dmd gene; Duchenne muscular dystrophy; dystrophin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 12446)
AUTHORS Rosenthal, A., Speer, A., Billwitz, H., Cross, G.S., Forrest, S.M. and Davies, K.E.
TITLE Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous
JOURNAL Nucleic Acids Res. 17 (13), 5391 (1989)
MEDLINE 89345106
PUBMED 266885
REFERENCE 2 (bases 1 to 12446)
AUTHORS Rosenthal, A.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1989) Rosenthal A., Akademie der Wissenschaften der DDR, Zentralinstitut fuer Molekularbiologie, Robert-Rössle Str.10, 1115 Berlin Buch, DDR
```

see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 456, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.

FEATURES

source
1. 12446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X chromosomal, Xp21."
/tissue_type="muscle"
/dev_stage="fetal and adult."
99. 11156
/note="dystrophin (AA 1 - 3685)"
/codon_start=1
/protein_id="CAA32479.1"
/db_xref="GI:30846"
/db_xref="SWISS-PROT:P11532"

CDS

translation="MLWEEVEDCYERDVQKFTFTKWNVAQPSKFGKHENLFSDL
QDGRLLDLEGLTGQKLPEKKGSTRVHALNPNKALRVLQNNVDLVNIGSTDIVDG

QY	822	TTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGACGACGACCTATTATTCGAGCGC	881
Db	8548	TTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGACGACGACCTATTATTCGAGCGC	8607
QY	882	ACTTTCAGCAGTTTCAGAAAGCAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAA	941
Db	8608	ACTTTCAGCAGTTTCAGAAAGCAACGATGTACATAGGGCCCTTCAAGAGGGAATTTGAAAA	8667
QY	942	CTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTGACGAATATTTCTGACAGAGCAGC	1001
Db	8668	CTAAAGAACCTGTAAATCATGAGTACTCTTTGAGACTGTGACGAATATTTCTGACAGAGCAGC	8727
QY	1002	CTTTTGAAGGACTAGAGAAACTCTACCAGAGCCACGAGAGCTGCCTCCTGAGGAGAG	1061
Db	8728	CTTTTGAAGGACTAGAGAAACTCTACCAGAGCCACGAGAGCTGCCTCCTGAGGAGAG	8787
QY	1062	CCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA	1121
Db	8788	CCCAAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAA	8847
QY	1122	AATTGAACTTCGACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCTCTGAAAGAATCC	1181
Db	8848	AATTGAACTTCGACTCCGCTGACTGCGCAGAGAAAAATAGATGAGACCTCTGAAAGAATCC	8907
QY	1182	AGGAACCTTCAAGAGGCCACGGATGAGCTCGAAGCTCAAGCTGCGCCAAAGCTGAGTGATCA	1241
Db	8908	AGGAACCTTCAAGAGGCCACGGATGAGCTCGAAGCTCAAGCTGCGCCAAAGCTGAGTGATCA	8967
QY	1242	AGGATCCTGGCAGCCGCTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACTCTCGAGA	1301
Db	8968	AGGATCCTGGCAGCCGCTGGCGGATCTCCTCATTTGACTCTCTCCAAAGATCACTCTCGAGA	9027
QY	1302	AAGTCAAGGCACTTCGAGAGAAATTTGCCCTCTGAAAGAGAAAGCTGAGCCACGCTCAATG	1361
Db	9028	AAGTCAAGGCACTTCGAGAGAAATTTGCCCTCTGAAAGAGAAAGCTGAGCCACGCTCAATG	9087
QY	1362	ACCTTGCTGGCAGCTTACACATTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACTC	1421
Db	9088	ACCTTGCTGGCAGCTTACACATTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACTC	9147
QY	1422	TGGAAGACCTTGAACACACAGATGGAAGCTTCTCGAGGTGGCCGTGAGGACCGAGTCAAGC	1481
Db	9148	TGGAAGACCTTGAACACACAGATGGAAGCTTCTCGAGGTGGCCGTGAGGACCGAGTCAAGC	9207
QY	1482	AGCTGCATGAAGCCACAGGGACTTTTGGTTCAGACTCTCAGACACTCTCTTTCCAGCTCTG	1541
Db	9208	AGCTGCATGAAGCCACAGGGACTTTTGGTTCAGACTCTCAGACACTCTCTTTCCAGCTCTG	9267
QY	1542	TCCAGGTCCTGGGAGAGAGCCATCTCGCAAAACAAAGTGCCTACTATATCAACCAACG	1601
Db	9268	TCCAGGTCCTGGGAGAGAGCCATCTCGCAAAACAAAGTGCCTACTATATCAACCAACG	9327
QY	1602	AGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTAGCTG	1661
Db	9328	AGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTAGCTG	9387
QY	1662	ACCTGAATAATGTGAGATCTTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCGAGA	1721
Db	9388	ACCTGAATAATGTGAGATCTTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCGAGA	944
QY	1722	AGGCCCTTTGTCTGGATCTCTTGAGCCCTGTCACTGTGATGCTTTGGACCAAGCACA	1781
Db	9448	AGGCCCTTTGTCTGGATCTCTTGAGCCCTGTCACTGTGATGCTTTGGACCAAGCACA	9507
QY	1782	ACCTCAAGCAAAATGACAGGCCATGGATATCTTCGAGATTAATTAATTTGTTGACCACTA	1841
Db	9508	ACCTCAAGCAAAATGACAGGCCATGGATATCTTCGAGATTAATTAATTTGTTGACCACTA	9567
QY	1842	TTTATGACCGCTGGAGCAAGAGCACAATAATTTGGTCAACGCTCCTCTCTCGTGGATA	1901
Db	9568	TTTATGACCGCTGGAGCAAGAGCACAATAATTTGGTCAACGCTCCTCTCTCGTGGATA	9627
QY	1902	TGTGTCTGAACTGGCTGCTGAAATGTTTATGATACGGGACGAAACAGGGAGGATCCGTGCTC	1961

NHKLTLGLTWNIIILEHWQYVNMKNIMAGLOFTNSEKILLSWVRQSTRNYPQWNINFT
 TWSGCLAINLIHHRPDLFDMSVWCQSAQRLHEAFNIARYQLGTEHLLDEPDY
 YITIPDKSLIMYIIPDPLPQVQVSEIAIQEVMLEPRPKVTKBEHFQHLHEQMHSYQ
 QITVSLAQGYERTSPKFRFSYVAQTAAYVITSDPTRSPVPSCHLEAPZKSGFSSL
 MSEVNLNRYQALBEVLSWLLSAEDTLAQAGEISNDVWVVKDQFTEHGYAMDLTAH
 QGRVNLQIGSKLGTGKLSEBETEVQEQMNLNRSWECILAVSWKXSNLHRLVM
 DLQNKGLKELNDWLTKTERTKMEPEPLGLEDLAKRQVQKHVQLQEDLOESQVRVN
 SILTHMVVVVDESSGHATAALEEQKVLGDRWANI CRWTEDRWVLIQDILLKQQRUTE
 QOLLSFAMISEXEDAVNKITQYKDNQEMLSLQKLAIVLAKDLKXKQSGFSEIYK
 EDDTLTKNKSVTQKTEAWLONFARCDWNLVQKLEKTAQISQAVTTTQFSLQITVTM
 EVTVTTTREQILVKAQEBELPPPKPKRQITVDSIRKRLDVIDTELHSMVITSEAE
 VLGSPFAPLFRXGHSFSLRQKELKREKVALEKXKPKLQDASRSQASQWVMEGNA
 DSTKQASQOLNRSWITFQCLLSERLNHLWLEYQNNIIAYFNLQQLQEQMTTAEWLKIO
 PTTFSPPTAKLSQKICXDEVNLSUGUQPIERLUKTIQSIALKKXGQPMFLDADVAF
 THNFQGVSDVQAQKEBTLITDPLPMRYQETMSAIRTWVQSESTKLSIPQSVVDY
 EINMEQRLQALQSLQOOSGSLYITLTKVMSKCAPSEISRKTYCSFSEIEGRWK
 KLSLOLVHCHQKLEQWKNLARKIQHIIQTLTKMAWVDVFLKEEWPALGDSIELLKQ
 KOCRLALDITQIPLSNVNEGGHQTKNBAEPFASRLRETKELTELKNTOWDHWCCQVY
 AKREALKGLSELTQSLQNDLSEMEHWMTAEZEYLERDPRYKTDLOELQAFEMKRAK
 EBAQXKAYKULLTSVNSVYTAQPPVAQALKELEPLITNYOMLCTRLNKCCKTLE
 EYVACWHELLTSQKXKMWLNNEFPLKTIENIPGGAEEISEVLDSIENLNRHSEDPN
 NQIRILAQTLTGGVWNLINELNEETLNSRWLEHAEVARRKQLSOSIQASCTENS
 LHLIQESLFTIDQKLAAYTADKPAQAPQBAKIQSDLTSLHSEIEMKNGNOKGEA
 AQVLSQIDVAKQLQGVSMKFRIFQKPAFNEQRLQESKMLDVEHVMYIPALLETALHY
 QVGVQSOLNHLKYLKSLSEVKEVEMVITKTRQIVQKOTENPKLDERVETALKLHY
 NELGAVPCHVKOOLKELKLSRQKREKMNVLTEMLATMELTKRSALVESPMNLDS
 VAWGAKTKEIKKQVHLKSI TEVGEALKTVLGCKEVLVEDKLSLINSNLIWATPSAE
 EWNLNLETKQHWNTQNDVHTKWIQIQTDLDESEKKKPOQKEVNLKRLKABELN
 IREKVDSTQDANLMAHGHCRKLVFQISLNLHPFAAISHRITKQASIPKLELE
 QFNSDIQKLELEAIQGVNLKEEDPKOMNEDNEGTVKRELLQRDLNQRITDER
 KSEIKI KOOLLQTKHNAIKOLRSRQKALIEI SHOWYKQRADDLKCLDIEKKL
 ASLPEPRDKRILKEIDRELVKBEELNVAARQESLGEDGAAMAVPTQISLKS KWRE
 IESKQAPRFLNAIQTDRVREBTVMNEMDPELISYVPTTYLTETHTVSOALLVEQ
 LNLNAPLCAKDPEDILFKQESLSANIKDSIQSSGRIDI IHSKKTALQASATPVERKL
 QEALSQDFTQEWKYNKQDGRFDRSEKVRKPHYDIIKFNQWLTAEQFLKRTQI
 PEWEHAKYKWLKELQDQIGQRTVTRTANATGEEIIQOSSSTASIQEKLGSLN
 RWEQVQKLSDRKLEBQKNTLSFQFDLNEFVILMEADNATSIPLSPGKQQLKE
 KLBOVQLWELPLRGOLIKUNETGPGVLVSAPISPEEQDKLENKLTQTNLQIKVS
 RALPEKQEAQDQIDGLEKLEBDELOLNHLLWLSPIRQOLEITYNQNOBQGF
 VKETEIVQKQEPDVEILSKQHLKKEKPATQPVKRLDOLSEKAWNRLLOELRA
 KQDPLAGLTTGASPTQVTLVTPQVTKETASIKLEMPSSLMLEVPALADFNRAWT
 ELTDWLSQDVQKQVQKNSLARSLEASDDOMKELHLSIQELLVWLKQDDEL
 NKTNSQEAQITIDTRERQNGDEBQHLQNRRLQKQEMLKQSTOWLEAKAEBOVL
 GQARAKLSEKGPYTVDAIQKTIETRELAQDKRWQNTVNDVALLKLFEDSA
 PUKWIMYINENASWRSIHKVSREKALBETHRLQOQFPLDEKFLAMLTEAFTA
 NVLQDATRKELLEDKSGVKELMKQWOLQGEIISAHTDVHNLDSQKILRLSEGD
 DAVLLQRLDNMKNWSELKSLARSLEASDDOMKELHLSIQELLVWLKQDDEL
 SRAPIGDGFPAQWQVHRAPKELKTKFVIMSTLTVTRIFETRQPLEGLEKLYQ
 EPELPEPBERAQNTRLLKQKAEVNTWEKLNHLSAQQKIDQWKTLELRLQELQARL
 ELDOLKRAEVINGQSPGVGDLILSDIQLHLEKVALRGAEIAPLENVSHVNDLQARL
 TLTGILSPVNLVSHETQVRLKQAVEDRYQLAHNRDPFQASQHEFSLTSVQGP
 WZRAISPNKPVYINETHQKQWQVTLQSLADLNNVNFSAYTAKMLRQAKA
 LCLDILSLSACDALQDHLNKQDPMIDIOINCLTITVYDRLOEHNHNVNPLGVD
 MCLNLLNVYDTQGRVILPDKPTGIIISLCAHLEDKARYLFPKQASVTFGCDORRL
 GLLLHDSIQIPRGVLEPFGSNIENIEPVSRCQFPAKKPIEAALFDNMRLPQSM
 WMLVPLVAAATCTTSGEDYAKLKNKFKRYFAKHPMWGLYPCQCSFSGSRVAGKH
 KMLPYHWEVETCTTSGEDYAKLKNKFKRYFAKHPMWGLYPCQCSFSGSRVAGKH
 PVTILNVPVDSAPSNQSPQLSHDDTHRIEHTYASLAEMENSNGSYLSINPESID
 DEHLIIQVYQASNLQSPSQRSQPAQIILSLESEGERELIADLEENRNLOABY
 DRUKQEHGKILRQLLPQPEAMPTQSPQSDRAELTAEAKLIRQKHGLEARMQILDEH
 NKLQESOLHRLQLLPQPEAKVNGTYSPPSTLSQSDSSQPMLLRVVGSQTSDSM
 GEEDLSPQDSTIGLEWNEQLNNSFFSSRRGRTFGKPMREDTM"

9628 TGAGTCTGAAGTGGCTGCTGAATGTTTATCATACGGAGCAACAGGAGGATCCGCTGCC 9687
1962 TGCTTTTAAAGCTGGCATCATTTCCCTCTGTAAGACACATTTGGAAGACAAGTACAGAT 2021
9688 TGCTTTTAAAGCTGGCATCATTTCCCTCTGTAAGACACATTTGGAAGACAAGTACAGAT 9747
2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGCCCTCC 2081
9748 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGCCCTCC 9807
2082 TTCTGCATGATTTCTATCCAA 2101
9808 TTCTGCATGATTTCTATCCAA 9827

RESULT 10
AX409637
LOCUS AX409637 13957 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2284 from Patent WO0229103.
ACCESSION AX409637
VERSION AX409637.1 GI:21442342
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 2284 11-APR-2002;
JOURNAL GENE LOGIC INC (US)
FEATURES
source
1. 13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M18533"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 63.2%; Score 1328.8; DB 6; Length 13957;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

762 GGGTCTTTTACAGACAGTCTGACCACTGGAGGCTGTGCACCTTCTCTGCAGGAC 821
8598 GGTCCCATTTTGAAGCCAGTCTGACCGTGGAAAGCTGTGCACCTTCTCTGCAGGAC 8657
822 TTCTGGTGTGGTACAGCTCAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGG 881
8658 TTCTGGTGTGGTACAGCTCAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGGG 8717
882 ACTTTCAGCAGTTTCAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAATA 941
8718 ACTTTCAGCAGTTTCAAGCAGACGATGTACATAGGCGCTTCAAGAGGGAATTGAATA 8777
942 CTAAGAACCTGTAACTCATCAGTACTCTTGAGACTGTACGAATTTCTGACAGACGAC 1001
8778 CTAAGAACCTGTAACTCATCAGTACTCTTGAGACTGTACGAATTTCTGACAGACGAC 8837
1002 CTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTCTGAGGAGAGAG 1061
8838 CTTTGGAGGACTAGAGAACTCTACAGAGGCCAGAGAGCTGCCTCTCTGAGGAGAGAG 8897
1062 CCGAGATGTCACTCGCTTCTAGGAAGCAGGCTGAGGAGTCAATCTAGTGGGAA 1121
8998 CCCAGATGTCACTCGCTTCTAGGAAGCAGGCTGAGGAGTCAATCTAGTGGGAA 9957
1122 AATTGAACCTCGACTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCC 1181
8998 AATTGAACCTCGACTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCC 9017

QY 1182 AGGAACCTTCAAGAGGCGCACCGATGAGCTGGAAGCTGAGCTGAGCTGAGCTGATCA 1241
DB 9018 AGGAACCTTCAAGAGGCGCACCGATGAGCTGGAAGCTGAGCTGAGCTGAGCTGATCA 9077
QY 1242 AGGATCCTGCGAGCGCGTGGCGATCTCTCTGATGACTCTCTCCTCAAGATCACTCGAG 1301
DB 9078 AGGATCCTGCGAGCGCGTGGCGATCTCTCTGATGACTCTCTCCTCAAGATCACTCGAG 9137
QY 1302 AAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAGAGAAACGCTGAGCAAGTCAATG 1361
DB 9138 AAGTCAAGGCACTTCGAGGAGAAATTCGCGCTCTGAAGAGAAACGCTGAGCAAGTCAATG 9197
QY 1362 ACCTTGTCTGCGAGCTTACCACTTTGGGCAATTCAGCTCTCAGCTATACCTCAGCACTC 1421
DB 9198 ACCTTGTCTGCGAGCTTACCACTTTGGGCAATTCAGCTCTCAGCTATACCTCAGCACTC 9257
QY 1422 TGAAGACCTGGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAAGC 1481
DB 9258 TGAAGACCTGGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAAGC 9317
QY 1482 AGTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTG 1541
DB 9318 AGTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTG 9377
QY 1542 TCCAGGCTCCCTGGAGAGAGCCATCTCGCAACAAAGTGCCTACTATATCAACACG 1601
DB 9378 TCCAGGCTCCCTGGAGAGAGCCATCTCGCAACAAAGTGCCTACTATATCAACACG 9437
QY 1602 AGACTCAACAACTTGTGGAGCAATCCCAAAATGACAGAGCTCTACAGCTTTAGCTG 1661
DB 9438 AGACTCAACAACTTGTGGAGCAATCCCAAAATGACAGAGCTCTACAGCTTTAGCTG 9497
QY 1662 ACCTGAATATGTGAGATTTCTGAGCCTGTGAGCCTGTGAGCCTGTGAGCCTGTGAGC 1721
DB 9498 ACCTGAATATGTGAGATTTCTGAGCCTGTGAGCCTGTGAGCCTGTGAGCCTGTGAG 9557
QY 1722 AGGCCCTTTGCTTGGATCTCTTGGAGCCTGTGAGCCTGTGAGCCTGTGAGCCTGTGAG 1781
DB 9558 AGGCCCTTTGCTTGGATCTCTTGGAGCCTGTGAGCCTGTGAGCCTGTGAGCCTGTGAG 9617
QY 1782 ACCTCAAGCAAAATGACAGCCATCTGAGTATCTGAGTATTTATTTGTTGACCACTA 1841
DB 9618 ACCTCAAGCAAAATGACAGCCATCTGAGTATCTGAGTATTTATTTGTTGACCACTA 9677
QY 1842 TTTATGACCGCTTGGAGCAAGACCAAACTTTTGGTCAACGCTCTCTCTGCTGGGATA 1901
DB 9678 TTTATGACCGCTTGGAGCAAGACCAAACTTTTGGTCAACGCTCTCTCTGCTGGGATA 9737
QY 1902 TGTGCTGAACCTGCTGCTGAAATGTTTATGATCGGACGAAACAGGAGGATCGGTGTC 1961
DB 9738 TGTGCTGAACCTGCTGCTGAAATGTTTATGATCGGACGAAACAGGAGGATCGGTGTC 9797
QY 1962 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTTGGAGACAAAGTACAGAT 2021
DB 9798 TGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTTGGAGACAAAGTACAGAT 9857
QY 2022 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGCCCTCC 2081
DB 9858 ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGCCCTCC 9917
QY 2082 TTCTGCATGATTTCTATCCAA 2101
DB 9918 TTCTGCATGATTTCTATCCAA 9937

RESULT 11
AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS Homo sapiens (human)
SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Chamberlain, J.S. and Harper, S.O.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES
source 1..13957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN

Query Match 63.2%; Score 1328.8; DB 6; Length 13957;
Best local Similarity 99.5%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 762 GGTTCTCTTTTACAGACAGTTCTGACAGTGGAGCGTCTGCACCTTTCTTCGAGGAAC 821
Db 8598 GGTCCCATTTGGAAGCAGTTCTGACAGTGGAGCGTCTGCACCTTTCTTCGAGGAAC 8657
QY 822 TTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGAGGCG 881
Db 8658 TTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGAGGCG 8717
QY 882 ACTTTCAGCAGTTTCAAGACGAGTGTACATAGGGCCCTCAAGAGGGGAATTTGAAAA 941
Db 8718 ACTTTCAGCAGTTTCAAGACGAGTGTACATAGGGCCCTCAAGAGGGGAATTTGAAAA 8777
QY 942 CTAAGAAGCCTGTAAATCATGAGTACTTCTGAGCTGTAGCAATATTTCTGACAGAGCAGC 1001
Db 8778 CTAAGAAGCCTGTAAATCATGAGTACTTCTGAGCTGTAGCAATATTTCTGACAGAGCAGC 8837
QY 1002 CTTTGGAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCGAGGAGAGAG 1061
Db 8838 CTTTGGAGGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCGAGGAGAGAG 8897
QY 1062 CCAGAAATGTCACTCGGCTTCTACGAAAGCAGCTGAGAGGTCAATATGAGTGGGAAA 1121
Db 8898 CCAGAAATGTCACTCGGCTTCTACGAAAGCAGCTGAGAGGTCAATATGAGTGGGAAA 8957
QY 1122 AATTGAACTTGCCTCGCTGCTGAGCAGAGAAATAGATGAGACCTTTGAAAGACTCC 1181
Db 8958 AATTGAACTTGCCTCGCTGCTGAGCAGAGAAATAGATGAGACCTTTGAAAGACTCC 9017
QY 1182 AGGAATCTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCA 1241
Db 9018 AGGAATCTCAAGAGGCCACGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCA 9077
QY 1242 AGGATCTTGGCAGCCCGTGGGCGATCTCTCTGACTCTCTCCAGATCACTCTCGAGA 1301
Db 9078 AGGATCTTGGCAGCCCGTGGGCGATCTCTCTGACTCTCTCCAGATCACTCTCGAGA 9137
QY 1302 AAGTCAAGGACATCTCGAGGAGAAATTCGGCTCTGAAAGAGAACGTCGAGCCAGTCAATG 1361
Db 9138 AAGTCAAGGACATCTCGAGGAGAAATTCGGCTCTGAAAGAGAACGTCGAGCCAGTCAATG 9197
QY 1362 ACCTTGTCTGCGAGCTTACACATTTTGGGCGATGAGTCTCAGCTTACCGTATACCTCAGCACTC 1421
Db 9198 ACCTTGTCTGCGAGCTTACACATTTTGGGCGATGAGTCTCAGCTTACCGTATACCTCAGCACTC 9257
QY 1422 TGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAAGC 1481
Db 9258 TGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAAGC 9317
QY 1482 AGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTG 1541
Db 9318 AGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTG 9377
QY 1542 TCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 1601

Db 9378 TCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 9437
QY 1602 AGACTCAAAACACTTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 1661
Db 9438 AGACTCAAAACACTTCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTG 9497
QY 1662 ACCTGAATATGTCAGATCTCAGCTTATAGAGTCCCAATGAACTCCGAAAGACTCCAGA 1721
Db 9498 ACCTGAATATGTCAGATCTCAGCTTATAGAGTCCCAATGAACTCCGAAAGACTCCAGA 9557
QY 1722 AGGCCCTTTGCTTGGATCTCTTGAGCTGTGAGCTGCATGTGATGCTTGGACCAAGCACA 1781
Db 9558 AGGCCCTTTGCTTGGATCTCTTGAGCTGTGAGCTGTGATGCTTGGACCAAGCACA 9617
QY 1782 ACCTCAAGCAAAATGACAGCCCATGATATCTGTCAGATTTAATTTGTTGACCACTA 1841
Db 9618 ACCTCAAGCAAAATGACAGCCCATGATATCTGTCAGATTTAATTTGTTGACCACTA 9677
QY 1842 TTTATGACCCCTGGGAGCAGCACAACAATTTGCTCAACGTCCTCTCTGCGTGGATA 1901
Db 9678 TTTATGACCCCTGGGAGCAGCACAACAATTTGCTCAACGTCCTCTCTGCGTGGATA 9737
QY 1902 TGTGCTCTGAACCTGGCTGCTGAATTTTATGATACGGGACCAACAGGAGAGTCCGTGTC 1961
Db 9738 TGTGCTCTGAACCTGGCTGCTGAATTTTATGATACGGGACCAACAGGAGAGTCCGTGTC 9797
QY 1962 TGTCTTTTAAACTGGCATTTTCCCTGTGTHAAGCACAATTTGGAGACAAGTACAGAT 2021
Db 9798 TGTCTTTTAAACTGGCATTTTCCCTGTGTHAAGCACAATTTGGAGACAAGTACAGAT 9857
QY 2022 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCAGCGTGGGCTCC 2081
Db 9858 ACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCAAGCAGCGTGGGCTCC 9917
QY 2082 TCTGCAATGATTTATCCAA 2101
Db 9918 TCTGCAATGATTTATCCAA 9937

RESULT 12
HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000
LOCUS Homo sapiens dystrophin (DMD) mRNA, complete cds.
DEFINITION ML8533 M17154 M18026 M20250
ACCESSION ML8533.1 GI:181856
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1699)
AUTHORS Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C. and Kunkel, L.M.
TITLE Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals
JOURNAL Cell 50 (3), 509-517 (1987)
MEDLINE 87273512
PUBMED 3607877
REFERENCE 2 (bases 1678 to 3830)
AUTHORS Hoffman, E.P., Monaco, A.P., Feener, C.C. and Kunkel, L.M.
TITLE Conservation of the Duchenne muscular dystrophy gene in mice and humans
JOURNAL Science 238 (4825), 347-350 (1987)
MEDLINE 88018015
PUBMED 3659917
REFERENCE 3 (bases 1 to 13957)
AUTHORS Koenig, M., Monaco, A.P. and Kunkel, L.M.
TITLE The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein
JOURNAL Cell 53 (2), 219-226 (1988)
MEDLINE 88194521

Db	9318	AGCTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGACCTTTCTTCCAGCTG	9377
Qy	1542	TCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG	1601
Db	9378	TCAGGGTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG	9437
Qy	1602	AGACTCAAAACAACTTGGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGCTTTTAGCTG	1661
Db	9438	AGACTCAAAACAACTTGGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGCTTTTAGCTG	9497
Qy	1662	ACCTGAATTAATGTAGATTTCTCAGCTTATAGAGCTGCCATGAATCCGAAAGCTGCAGA	1721
Db	9498	ACCTGAATTAATGTAGATTTCTCAGCTTATAGAGCTGCCATGAATCCGAAAGCTGCAGA	9557
Qy	1722	AGGCCCTTTGGTGGATCTCTTGGAGCTCTCAGCTGCATGTGATGCTTGGACCAACACA	1781
Db	9558	AGGCCCTTTGGTGGATCTCTTGGAGCTCTCAGCTGCATGTGATGCTTGGACCAACACA	9617
Qy	1782	ACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTAATTAATTTGACCACTA	1841
Db	9618	ACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTAATTAATTTGACCACTA	9677
Qy	1842	TTTATGACCGCTGGAGCAAGCAACAAATTTGGTCAACAGCTCCCTCTCGGTGGATA	1901
Db	9678	TTTATGACCGCTGGAGCAAGCAACAAATTTGGTCAACAGCTCCCTCTCGGTGGATA	9737
Qy	1902	TGTGTCTGAACCTGGCTGCTGAATTTGATATAGAGCAAGCAAGGAGGATCCGTGTCC	1961
Db	9738	TGTGTCTGAACCTGGCTGCTGAATTTGATATAGAGCAAGCAAGGAGGATCCGTGTCC	9797
Qy	1962	TGTGTCTGAACCTGGCTGCTGAATTTGATATAGAGCAAGCAAGGAGGATCCGTGTCC	2021
Db	9798	TGTGTCTGAACCTGGCTGCTGAATTTGATATAGAGCAAGCAAGGAGGATCCGTGTCC	9857
Qy	2022	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGCTGGGCTCC	2081
Db	9858	ACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGCTGGGCTCC	9917
Qy	2082	TTCTGCATGATTTCTATCCAA 2101	
Db	9918	TTCTGCATGATTTCTATCCAA 9937	
RESULT 13			
LOCUS	AR220819	13977 bp	DNA
DEFINITION	AR220819	Sequence 60 from patent US 6426186.	linear
ACCESSION	AR220819		
VERSION	AR220819.1	GI:23327696	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 13977)		
AUTHORS	Jones K.A., Volkmuth, W. and Walker M.G.		
TITLE	Bone remodeling genes		
JOURNAL	Patent: US 6426186-A 60 30-JUL-2002;		
FEATURES	Location/Qualifiers		
source	1..13977		
BASE COUNT	4596 a 2765 c 3120 g 3453 t		43 others
ORIGIN			
Query Match	62.7%	Score 1317.8;	DB 6; Length 13977;
Best Local Similarity	99.4%	Pred. No. 0;	
Matches 1333;	Conservative 0;	Mismatches 7;	Indels 1; Gaps 1;
Qy	762	GGGTTCTTTTCAAGACAGTTCGACAGTGGAGAGCGTTCGACCTTCTCTCGAGGAAAC	821
Db	8598	GGTCCCATTTGGAAGCCAGTTCTGACCAAGTGGAGAGCGTTCGACCTTCTCTCGAGGAAAC	8657
Qy	822	TTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGGACCACTATTGGAGCG	881

Db	8658	TTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGAGGACCACTATTGGAGCG	8717
Qy	882	ACTTTCCAGCAGTTTCAAGACAGAAACGATGTACATAGGCTTTCAAGAGGAAATTGA	941
Db	8718	ACTTTCCAGCAGTTTCAAGACAGAAACGATGTACATAGGCTTTCAAGAGGAAATTGA	8777
Qy	942	CTAAAGAACTGTATCATGAGTACTCTTTGAGACTGTAGCAATATTTCTGACAGAGCAGC	1001
Db	8778	CTAAAGAACTGTATCATGAGTACTCTTTGAGACTGTAGCAATATTTCTGACAGAGCAGC	8837
Qy	1002	CTTTGAGAGACTTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAG	1061
Db	8838	CTTTGAGAGACTTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAG	8897
Qy	1062	CCAGAAATGTACTCGGCTTCTACGAAAGAGCGCTGAGAGGTCAATAGTGTGGGAA	1121
Db	8898	CCAGAAATGTACTCGGCTTCTACGAAAGAGCGCTGAGAGGTCAATAGTGTGGGAA	8957
Qy	1122	AATTGAACTGCACTCGGCTGACTGCGCAGAGAAATAGATGAGACCTTTGAAAGACTCC	1181
Db	8958	AATTGAACTGCACTCGGCTGACTGCGCAGAGAAATAGATGAGACCTTTGAAAGACTCC	9017
Qy	1182	AGGAATTTCAAGAGGCCAGGATGAGTGGACCTCAAGCTGGCCAAAGCTGAGGTATCA	1241
Db	9018	AGGAATTTCAAGAGGCCAGGATGAGTGGACCTCAAGCTGGCCAAAGCTGAGGTATCA	9077
Qy	1242	AGGATCTCTGGCAGCCGCTGGGCGATCTCTCTTCAAGATCACTCTCGAGA	1301
Db	9078	AGGATCTCTGGCAGCCGCTGGGCGATCTCTCTTCAAGATCACTCTCGAGA	9137
Qy	1302	AAGTCAAGGCACTTCGAGGAGAAATTTGGCTCTGAAAGAGAAAGCTGAGCCAGTCAATG	1361
Db	9138	AAGTCAAGGCACTTCGAGGAGAAATTTGGCTCTGAAAGAGAAAGCTGAGCCAGTCAATG	9197
Qy	1362	ACCTTGCTGCGAGCTTACCACTTTGGGCACTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT	1421
Db	9198	ACCTTGCTGCGAGCTTACCACTTTGGGCACTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT	9257
Qy	1422	TGGAAGACCTTGAAACCAAGAGCTTCTGAGGTTGGCTCGAGGAGCGGCTCGAGGACCGAGT	1481
Db	9258	TGGAAGACCTTGAAACCAAGAGCTTCTGAGGTTGGCTCGAGGAGCGGCTCGAGGACCGAGT	9317
Qy	1482	AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTCTTCTTCCACGCTG	1541
Db	9318	AGCTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTCTTCTTCCACGCTG	9377
Qy	1542	TCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACAG	1601
Db	9378	TCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACAG	9437
Qy	1602	AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCACTTTAGCTG	1661
Db	9438	AGACTCAAAACAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCACTTTAGCTG	9497
Qy	1662	ACCTGAATTAATGTAGATTTCTCAGCTTATAGGATGCGCATGAAACTCCGAAAGCTGCAGA	1721
Db	9498	ACCTGAATTAATGTAGATTTCTCAGCTTATAGGATGCGCATGAAACTCCGAAAGCTGCAGA	9557
Qy	1722	AGGCCCTTTGGTGGATCTCTTGGAGCTGTGATGATGATGATGATGATGATGATGATGATGAT	1781
Db	9558	AGGCCCTTTGGTGGATCTCTTGGAGCTGTGATGATGATGATGATGATGATGATGATGATGAT	9617
Qy	1782	ACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTAATTAATTTGACCACTA	1841
Db	9618	ACCTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTAATTAATTTGACCACTA	9677
Qy	1842	TTTATGACCGCTGGAGCAAGCAACAAATTTGGTCAACAGCTCCCTCTCGGTGGATA	1901
Db	9678	TTTATGACCGCTGGAGCAAGCAACAAATTTGGTCAACAGCTCCCTCTCGGTGGATA	9737
Qy	1902	TGTGTCTGAACCTGGCTGCTGAATTTGATATAGAGCAAGCAAGGAGGATCCGTGTCC	1961
Db	9738	TGTGTCTGAACCTGGCTGCTGAATTTGATATAGAGCAAGCAAGGAGGATCCGTGTCC	9797

```
QY 1962 TGCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAGAT 2021
Db 9798 TGCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTACAGAT 9857
QY 2022 ACCTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCGAGGCT-GGGCCCTC 2080
Db 9858 ACCTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGTGACCAAGCGAGGCTGGGGCCCTC 9917
QY 2081 CTTCGTGATGATTCATCCAA 2101
Db 9918 CTTCGTGATGATTCATCCAA 9938

RESULT 14
LOCUS E30219 4402 bp DNA linear PAT 18-JUN-2001
DEFINITION Shortened dystrophin.
ACCESSION E30219
VERSION E30219.1 GI:13017026
KEYWORDS JP 1999318467-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4402)
AUTHORS Sinichi, I.
TITLE Shortened dystrophin
JOURNAL Patent: JP 1999318467-A 2 24-NOV-1999;
SCIENCE & TECH AGENCY, NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY
COMMENT OS Unidentified
PN JP 1999318467-A/2
PD 24-NOV-1999
PR 08-MAY-1998 JP 1998142134
PI SINICHI TAKEDA
PC C12N15/09, A61K48/00, C12N15/00
CC Strandedness: Both;
CC Topology: Linear;
FH Key location/Qualifiers
FT source 1..4402
FT /organism='Unidentified'.
FEATURES
source
location/Qualifiers
1..4402
/organism='unidentified'
/mol type='genomic DNA'
/db xref='taxon:32644'

BASE COUNT 1329 a 1000 c 1019 g 1054 t
ORIGIN

Query Match 62.7%; Score 1317; DB 6; Length 4402;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 0; Indels 387; Gaps 1;

QY 1 GAGTATGCCCTACACACAGGCTGTTATGTATGTCACACCTCTGACCCCTACACGGAGCCCAT 60
Db 1095 GAGTATGCCCTACACACAGGCTGTTATGTATGTCACACCTCTGACCCCTACACGGAGCCCAT 1154
QY 61 TCCTTCACAGATTTGGAAGCTCCTGAAGCAAGTCAATTTGGCAGTTTCATTGATGAGAG 120
Db 1155 TCCTTCACAGATTTGGAAGCTCCTGAAGCAAGTCAATTTGGCAGTTTCATTGATGAGAG 1214
QY 121 TGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAGAAAGTATTATCGTGGCTTTTTC 180
Db 1215 TGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAGAAAGTATTATCGTGGCTTTTTC 1274
QY 181 TGCTGAGACACATTTGACGACCAAGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA 240
Db 1275 TGCTGAGACACATTTGACGACCAAGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA 1334
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1335 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1394
```

```
QY 301 TAATATTCTCAATTGGGAAGTAAGTCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 360
Db 1395 TAATATTCTCAATTGGGAAGTAAGTCTGATTGGAAACAGGAAAATTTATCAGAAGATGAAGA 1454
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGTGGGAATGCCTCAGGCTAGC 420
Db 1455 AACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGTGGGAATGCCTCAGGCTAGC 1514
QY 421 TAGCATGAAAAAACAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT 480
Db 1515 TAGCATGAAAAAACAAGCAATTTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT 1574
QY 481 GAAAGAGTTGAATGACTGGCTTAACAAAAAACHAGAAGAAAGAACAGAAATGGAGGAAGA 540
Db 1575 GAAAGAGTTGAATGACTGGCTTAACAAAAAACHAGAAGAAAGAACAGAAATGGAGGAAGA 1634
QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAGGTGCTTCA 600
Db 1635 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACAATAGGTGCTTCA 1694
QY 601 AGAAGATCTAGAACAGAAACAAAGTCAGGCTCAATTTCTACTCATCTGTTGGTGGTAGT 660
Db 1695 AGAAGATCTAGAACAGAAACAAAGTCAGGCTCAATTTCTACTCATCTGTTGGTGGTAGT 1754
QY 661 TGATGAATCTAGTGSAGATCACGCAACTGCTGTTTGGAAAGAACAACTTAAGGTATTGGG 720
Db 1755 TGATGAATCTAGTGSAGATCACGCAACTGCTGTTTGGAAAGAACAACTTAAGGTATTGGG 1804
QY 721 AGATCGATGGSCNAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACAG 780
Db 1805 AGATCGATGGSCNAACATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACAG 1804
QY 781 TTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTTGGCTACAGCT 840
Db 1805 TTCTGACCAGTGGAAAGCGTCTGCACCTTTCTCTGCAGGAACCTTCTGTTGGCTACAGCT 1804
QY 841 GAAAGATGATGAATTAAGCCGCGCAGCACCTTATTGGAGCGGACTTTCCAGCAGTTTCAGAA 900
Db 1805 GAAAGATGATGAATTAAGCCGCGCAGCACCTTATTGGAGCGGACTTTCCAGCAGTTTCAGAA 1804
QY 901 CGAGAACGATGTACATAGGGGCTTCAAGAGGGGAATTTGAAAACCTGTAATCAT 960
Db 1805 CGAGAACGATGTACATAGGGGCTTCAAGAGGGGAATTTGAAAACCTGTAATCAT 1804
QY 961 GAGTACTCTTGAGACTGTACGAATATTCTTACAGAGCAGCCCTTTGGAAGGACTTAGAGAA 1020
Db 1805 GAGTACTCTTGAGACTGTACGAATATTCTTACAGAGCAGCCCTTTGGAAGGACTTAGAGAA 1804
QY 1021 ACTTTACCAGAGCCCACAGAGAGTGCCTCTCTGAGGAGAGAGCCACAGATGTCACTCGGCT 1080
Db 1805 ACTTTACCAGAGCCCACAGAGAGTGCCTCTCTGAGGAGAGAGCCACAGATGTCACTCGGCT 1804
QY 1081 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTTGCACTCCGC 1140
Db 1805 TCTACGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAAAAATTGAACCTTGCACTCCGC 1847
QY 1141 TCAGTCCAGAGAAAAATAGATGAGACCTTTGAAGACTCCAGGAATTCAGAGGCGCAC 1200
Db 1848 TCAGTCCAGAGAAAAATAGATGAGACCTTTGAAGACTCCAGGAATTCAGAGGCGCAC 1907
QY 1201 GATGAGTGCACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1260
Db 1908 GATGAGTGCACCTCAAGCTCGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGT 1967
QY 1261 GGGCGATCTCTCATTTGACTCTCTCCAAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGG 1320
Db 1968 GGGCGATCTCTCATTTGACTCTCTCCAAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGG 2027
QY 1321 AGAATTTCCGCTCTGAAGAGAACGTCGAGCCAGCTCAATGACCTTCTCGCCAGCTTAC 1380
Db 2028 AGAATTTCCGCTCTGAAGAGAACGTCGAGCCAGCTCAATGACCTTCTCGCCAGCTTAC 2087
QY 1381 CACTTTGGGCATTTCAGCTCTCACCGTATTAACCTCAGCACCTCTGGAAGACCTTGAAACACAG 1440
```


Db	2088	CACCTTGGGCAATCAGCTCTACCGTATACCTGAGCACTCTGGAGACCTGACACAG	2147
QY	1441	ATGGAAGCTTCTCAGGTGGCCGTGAGGACCGAGTCAAGGAGTGCATGAGGCCACAG	1500
Db	2148	ATGGAAGCTTCTCAGGTGGCCGTGAGGACCGAGTCAAGGAGTGCATGAGGCCACAG	2207
QY	1501	GGACTTTGGTCCAGGATCTCAGACATCTTCTTCCAGCTGTCCAGGTCCTGGGAG	1560
Db	2208	GGACTTTGGTCCAGGATCTCAGACATCTTCTTCCAGCTGTCCAGGTCCTGGGAG	2267
QY	1561	AGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAAAACCTTGCTG	1620
Db	2268	AGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAAAACCTTGCTG	2327
QY	1621	GGACCATCCAAATACAGAGCTTACAGCTCTTTAGCTGACCTGATATGTCAGATT	1680
Db	2328	GGACCATCCAAATACAGAGCTTACAGCTCTTTAGCTGACCTGATATGTCAGATT	2387
QY	1681	CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTGGATCT	1740
Db	2388	CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAAGAGGCCCTTTGCTGGATCT	2447
QY	1741	CTTGAGCCTGTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG	1800
Db	2448	CTTGAGCCTGTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG	2507
QY	1801	GCCCATGGATATCTCGAGATTAATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA	1860
Db	2508	GCCCATGGATATCTCGAGATTAATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA	2567
QY	1861	AGAGCAACAAATTTGGTCAACCTCCCTCTCTGCGTGGATATGTTGCTGAACTGGCT	1920
Db	2568	AGAGCAACAAATTTGGTCAACCTCCCTCTCTGCGTGGATATGTTGCTGAACTGGCT	2627
QY	1921	GAATGTTTATGATACGGAGCAAGAGGAGTCCCTGCTCTCTTTTAAACTGGCAT	1980
Db	2628	GAATGTTTATGATACGGAGCAAGAGGAGTCCCTGCTCTCTTTTAAACTGGCAT	2687
QY	1981	CATTTCCTCTGTAAGACCAATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGC	2040
Db	2688	CATTTCCTCTGTAAGACCAATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGC	2747
QY	2041	AGTTTCAACAGATTTTGTGACCAAGCGGCTGGGCTCTCTGCTGATGTTTATCCA	2100
Db	2748	AGTTTCAACAGATTTTGTGACCAAGCGGCTGGGCTCTCTGCTGATGTTTATCCA	2807
QY	2101	A 2101	
Db	2808	A 2808	
RESULT 15			
LOCUS	E30220	4402 bp	DNA linear PAT 18-JUN-2001
DEFINITION	Shortened dystrophin.		
ACCESSION	E30220		
VERSION	E30220.1		GI:13017027
KEYWORDS	JP 199318467-A/3.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 4402)		
AUTHORS	Sinichi T.		
TITLE	Shortened dystrophin		
JOURNAL	Patent: JP 199318467-A 3 24-NOV-1999;		
COMMENT	SCIENCE & TECH AGENCY,NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY		
	OS Unidentified		
	PN JP 199318467-A/3		
	PD 24-NOV-1999		
	PF 08-MAY-1998 JP 1998142134		
	PR		
	PI SINICHI TAKEDA		

PC	C12N15/09,A61K48/00,C12N15/00	
CC	Strandedness: Both;	
CC	Topology: Linear;	
EH	Key	Location/Qualifiers
FT	source	1..4402
FT	source	/Organism='Unidentified'.
FEATURES	Location/Qualifiers	
source	1..4402	
	/organism="unidentified"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32644"	
BASE COUNT	1339 a -984 c 1010 g 1069 t	
ORIGIN		
Query Match	57.0%; Score 1196.6; DB 6; Length 4402;	
Best Local Similarity	78.3%; Pred. No. 1.3e-289;	
Matches 1645; Conservative	0; Mismatches 69; Indels 387; Gaps 2;	
QY	1	GAGCTATGCTACACACAGCGTCTTATGTCACCACTCTGACCCCTACACGAGCCCAAT 60
Db	1095	GAGCTATGCTACACACAGCGTCTTATGTCACCACTCTGACCCCTACACGAGCCCAAT 1154
QY	61	TCCTTACACAGCATTTGGAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTCAATGATGAGAG 120
Db	1155	TCCTTACACAGCATTTGGAAGCTCTCTGAAGACAAGTCAATTTGGCAGTTCAATGATGAGAG 1214
QY	121	TGAAGTAACTGACCGCTTATCAACAGCTTTAGAGAGTATTATCTGTCGCTTCTTTC 180
Db	1215	TGAAGTAACTGACCGCTTATCAACAGCTTTAGAGAGTATTATCTGTCGCTTCTTTC 1274
QY	181	TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA 240
Db	1275	TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA 1334
QY	241	CCAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGGGTTGG 300
Db	1335	CCAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGGGTTGG 1394
QY	301	TAATATTCTCAATTTGGGAAGTAAAGCTGATTGGAACAGGAAATATTATCAAGATGAAGA 360
Db	1395	TAATATTCTCAATTTGGGAAGTAAAGCTGATTGGAACAGGAAATATTATCAAGATGAAGA 1454
QY	361	AACGGAAGTCAAGAGCAGATGAATCTCTTAATTTCAAGTGGGAATGCTCAGGTTAGC 420
Db	1455	AACGGAAGTCAAGAGCAGATGAATCTCTTAATTTCAAGTGGGAATGCTCAGGTTAGC 1514
QY	421	TAGCATGGAAGAACAGAGCAATTTACATAGATTTTAAATGATCTCCAGAAATCAGAACT 480
Db	1515	TAGCATGGAAGAACAGAGCAATTTACATAGATTTTAAATGATCTCCAGAAATCAGAACT 1574
QY	481	GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAGGAAATGAGGGAAGA 540
Db	1575	GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAGGAAATGAGGGAAGA 1634
QY	541	GCCTCTTGGACCTGATCTTGAGGACCTTAAAGCCCAAGTACAAACATAAGGTGCTTCA 600
Db	1635	GCCTCTTGGACCTGATCTTGAGGACCTTAAAGCCCAAGTACAAACATAAGGTGCTTCA 1694
QY	601	AGAAGATCTAGAACAAAGCAAGTCAAGTCAAGTCAATTTCTCACTCACATGTTGGTGTAGT 660
Db	1695	AGAAGATCTAGAACAAAGCAAGTCAAGTCAAGTCAATTTCTCACTCACATGTTGGTGTAGT 1754
QY	661	TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAGCAACTTAAGGTATTGGG 720
Db	1755	TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAGCAACTTAAGGTATTGGG 1814
QY	721	AGATCGATGGCAACATCTCTAGATGGAACAGAGCCGTGGGTCTTTTACAGACAG 780
Db	1815	AGATCGATGGCAACATCTCTAGATGGAACAGAGCCGTGGGTCTTTTACAGACAT 1874
QY	781	TTCTGACCAAGTGAAGGCTGTGACCTTTCTCTGACAGAACTTCTGTTGGCTACAGCT 840
Db	1875	CTTCTCAATGGCAACGCTCTTACTGAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 1934

QY 841 GAAAGATGATGAATTAAGCGGAGGACCACTATTGGAGGAGACTTTCCAGCAGTTTCAGAA 900
| | | | |
Db 1935 AAAGAAGATGCACT----- 1949
QY 901 GCAGAACGATGTACATAGGCGCTTCAAGAGGGAAATTGAAGAACTTAAGAACCTGTATATCAT 960
| | | | |
Db 1950 ----- 1949
QY 961 GAGTACTCTTTGAGACTGTACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA 1020
| | | | |
Db 1950 -----GAACRAGATTACACACACTGGCTTTAAAGATCAAAATGAA 1989
QY 1021 ACTCTACGAGAGCCGAGAGAGTGCCTCTCTGAGAGAGAGCCGAGATGTCACTCGGCT 1080
| | | | |
Db 1990 ATGTTATCAAG----- 2000
QY 1081 TCTACGAAACGAGCTCAGGAGGTCAATACTAGTGGGAAABAAATTGAACCTGCACCTCCGC 1140
| | | | |
Db 2001 ----- 2000
QY 1141 TGACTGCGAGAGAAATAAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGGCCAC 1200
| | | | |
Db 2001 ----- 2000
QY 1201 GGATGAGCTGGACCTCAAGCTGCGCCAAAGCTGAGTGAATCAAGGGATCTTGGCAGCCCGT 1260
| | | | |
Db 2001 ----- 2000
QY 1261 GGGCGATCTCCTCATTGACTCTCTCCAAGATCACCTCGAAGAGTCAAGGCACCTTCGAGG 1320
| | | | |
Db 2001 -----TCTCGAGAAAGTCAAGGCACCTTCGAGG 2027
QY 1381 CACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAG 1440
| | | | |
Db 2088 CACTTTGGGCAATTCAGCTCTCACCGTATACCTCAGCACTCTGGAAGACCTGAACACCAG 2147
QY 1441 ATGGAAGCTTTGAGAGTGCGCTCTGAGGAGACCGAGTGAAGCAGCTGCATGAAGCCACAG 1500
| | | | |
Db 2148 ATGGAAGCTTTGAGAGTGCGCTCTGAGGAGACCGAGTGAAGCAGCTGCATGAAGCCACAG 2207
QY 1501 GGACTTTGGTCCAGCACTCTCAGCACTTTCTTCCAGCTCTGTCAGGCTCCCTGGGAGAG 1560
| | | | |
Db 2208 GGACTTTGGTCCAGCACTCTCAGCACTTTCTTCCAGCTCTGTCAGGCTCCCTGGGAGAG 2267
QY 1561 AGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACACGAGACTCAAAACAACTTTGCTG 1620
| | | | |
Db 2268 AGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACACGAGACTCAAAACAACTTTGCTG 2327
QY 1621 GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTTAGCTGACCTGAATAATGTCAGATT 1680
| | | | |
Db 2328 GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTTAGCTGACCTGAATAATGTCAGATT 2387
QY 1681 CTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCGAGAGGCCCTTTGCTTGGATCT 1740
| | | | |
Db 2388 CTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCGAGAGGCCCTTTGCTTGGATCT 2447
QY 1741 CTTGAGCCTGTGACGTGCATGTGATGCTTGGACCGAGCAACCTCAGCAAAATGACCA 1800
| | | | |
Db 2448 CTTGAGCCTGTGACGTGCATGTGATGCTTGGACCGAGCAACCTCAGCAAAATGACCA 2507
QY 1801 GCCCATGGATATCTCGAGATTATTAATGTTTGAACCACTATTATGACCGCCTGGAGCA 1860
| | | | |
Db 2508 GCCCATGGATATCTCGAGATTATTAATGTTTGAACCACTATTATGACCGCCTGGAGCA 2567
QY 1861 AGAGCAACAAATTTGGTCAACGCTCCCTCTCTCGGTGGATATGTGCTGAACCTGGCTGCT 1920
| | | | |
Db 2568 AGAGCAACAAATTTGGTCAACGCTCCCTCTCTCGGTGGATATGTGCTGAACCTGGCTGCT 2627

QY 1921 GAATGTTTATGATACGGACGACGAAACAGGAGGATCCGCTGCTCTCTTTTAAACTGGCAT 1980
| | | | |
Db 2628 GAATGTTTATGATACGGACGACGAAACAGGAGGATCCGCTGCTCTCTTTTAAACTGGCAT 2687
QY 1981 CATTTCCTCTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
| | | | |
Db 2688 CATTTCCTCTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2747
QY 2041 AAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCTCTTCTGCATGATTTATATCCA 2100
| | | | |
Db 2748 AAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCTCTTCTGCATGATTTATATCCA 2807
QY 2101 A 2101
Db 2808 A 2808

Search completed: February 1, 2004, 23:51:21
Job time : 7497.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DN nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 499.045 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-12_COPY_900_3000

Perfect score: 2101

Sequence: 1 gagctatgctacacagg.....ttctgcattgattctatccaa 2101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2101	100.0	3510	24	Human dystrophin m
2	2101	100.0	4476	24	Adeno-associated v
3	1760.4	83.8	5339	24	DNA encoding mini-
4	1750.8	83.3	3531	24	Human dystrophin m
5	1750.8	83.3	4498	24	Adeno-associated v
6	1743	83.0	3858	24	Human dystrophin m
7	1743	83.0	4825	24	Adeno-associated v
8	1743	83.0	4848	24	Adeno-associated v

9	1743	83.0	5060	24	AAD37264	Adeno-associated v
10	1641	78.1	4414	24	AAD37260	Adeno-associated v
11	1629	77.5	3446	24	AAD37242	Human dystrophin m
12	1629	77.5	5462	24	ABX81999	DNA encoding mini-
13	1611.6	76.7	5417	24	ABX81997	DNA encoding mini-
14	1602	76.2	3999	24	AAD37234	Human dystrophin m
15	1602	76.2	4966	24	AAD37256	Adeno-associated v
16	1602	76.2	4990	24	AAD37262	Adeno-associated v
17	1419	67.5	4182	24	AAD37230	Human dystrophin m
18	1419	67.5	5149	24	AAD37255	Adeno-associated v
19	1328.8	63.2	2169	24	AAD37232	Human dystrophin m
20	1328.8	63.2	5952	22	AAD06794	DNA encoding mini-
21	1328.8	63.2	8689	24	ABX82000	Human dystrophin p
22	1328.8	63.2	11058	24	AAD37229	CDNA encoding huma
23	1328.8	63.2	11241	24	ABX82005	DNA encoding mini-
24	1328.8	63.2	11443	24	ABX82002	Sequence of human
25	1328.8	63.2	12923	10	AAH90338	Human breast cance
26	1328.8	63.2	13957	24	ABT10904	Human dystrophin g
27	1328.8	63.2	13957	24	ABX69900	Human dystrophin g
28	1328.8	63.2	13957	24	ABX95766	Gene #2284 used to
29	1328.8	63.2	13957	24	ABX81959	CDNA encoding huma
30	1324	63.0	1821	24	AAD37241	Human dystrophin r
31	1317.8	62.7	13977	24	ABX70403	Human bone remodel
32	1317	62.7	4402	21	AZ48567	A rod shortened dy
33	1196.6	57.0	4402	21	AZ48568	A rod shortened dy
34	1149.6	54.7	13815	24	ABX81960	CDNA encoding mous
35	1149.6	54.7	13815	24	ABT99799	Mouse ischaemic co
36	1149.6	54.7	19307	17	AAT27558	Shuttle vector pAd
37	1148	54.6	13815	19	AAV18895	Mus musculus cystr
38	937	44.6	1434	24	AAD37243	Human dystrophin r
39	785.4	37.4	1991	24	AAD37231	Human dystrophin N
40	777	37.0	1667	24	AAD37235	Human dystrophin N
41	718.8	34.2	4075	21	AZ48569	Human dystrophin N
42	667.2	31.8	3747	21	AAZ48566	A rod shortened dy
43	618.6	29.4	3275	10	AAH97129	Partial sequence o
44	613.6	29.2	3163	21	AAZ48571	A rod shortened dy
45	554.2	26.4	11096	24	ABX81962	CDNA encoding mous

ALIGNMENTS

RESULT 1

AAD37240
ID AAD37240 standard; DNA; 3510 BP.

XX AAD37240;

AC AAD37240;

DT 21-AUG-2002 (first entry)

XX Human dystrophin minigene delta3510.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

XX Homo sapiens.

XX WO200183695-A2.

PD 08-NOV-2001.

PF 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777F.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

PT New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -

PS Example 1; Page 51-52; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 other;

Query Match 100.0%; Score 2101; DB 24; Length 3510;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAGCTATGCTACACACAGGCTCTTATGTCCACCACTCTGACCCCTACACGAGGCCATT	60
DB	900	GAGCTATGCTACACACAGGCTCTTATGTCCACCACTCTGACCCCTACACGAGGCCATT	959
QY	61	TCCTTCAAGCAATTTGAAGCTCTCTGAAGCAAGTCATTGGCAGTTCAATGGAGAG	120
DB	960	TCCTTCAAGCAATTTGAAGCTCTCTGAAGCAAGTCATTGGCAGTTCAATGGAGAG	1019
QY	121	TGAAGTAACTCGACCGTTATCAACAGCTTTAGAGAGATTTATCTGCTGCTCTTTTC	180
DB	1020	TGAAGTAACTCGACCGTTATCAACAGCTTTAGAGAGATTTATCTGCTGCTCTTTTC	1079
QY	181	TGCTAGGACACATTCGACCAAGAGAGAGATTTCTAATGATGTGGAGAGTGGTGAAGA	240
DB	1080	TGCTAGGACACATTCGACCAAGAGAGAGATTTCTAATGATGTGGAGAGTGGTGAAGA	1139
QY	241	CCAGTTTCATACATGAGGGGTACATGATGGATTTGACAGCCATCAGGCGCGGTTGG	300
DB	1140	CCAGTTTCATACATGAGGGGTACATGATGGATTTGACAGCCATCAGGCGCGGTTGG	1199
QY	301	TAAATTTCTACAATTTGGAGTAAAGTGAATTTGAACACAGGAAATTTATCAGAAGATGAAGA	360
DB	1200	TAAATTTCTACAATTTGGAGTAAAGTGAATTTGAACACAGGAAATTTATCAGAAGATGAAGA	1259
QY	361	AACGTAAGTACAGAGCAGATGAATCTCCCTAAATTCAGATGGGAATGCTCAGGGTAGC	420
DB	1260	AACGTAAGTACAGAGCAGATGAATCTCCCTAAATTCAGATGGGAATGCTCAGGGTAGC	1319
QY	421	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT	480
DB	1320	TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAAACT	1379
QY	481	GAAGAGTTGAATCACTGGCTTAACAAAAACAGAGAAAGAACAGGAAATTTGAGGAGAGA	540
DB	1380	GAAGAGTTGAATCACTGGCTTAACAAAAACAGAGAAAGAACAGGAAATTTGAGGAGAGA	1439
QY	541	GCCTCTTGGACCTCATCTTGAAGACCTTAAAGCGCAAGTACAAACATAGGTTGCTTCA	600
DB	1440	GCCTCTTGGACCTCATCTTGAAGACCTTAAAGCGCAAGTACAAACATAGGTTGCTTCA	1499
QY	601	AGAAGATCTAGAACAAGAACAGTTCAGGGTCAATTTCTCTCACTCAATGTTGGTGTAGT	660
DB	1500	AGAAGATCTAGAACAAGAACAGTTCAGGGTCAATTTCTCTCACTCAATGTTGGTGTAGT	1559
QY	661	TGATGAATCTAGTGAGATCAAGCAATCTGCTGCTTTGGAAGAACAACTTAAAGTATTGGG	720
DB	1560	TGATGAATCTAGTGAGATCAAGCAATCTGCTGCTTTTGGAAAGAACAACTTAAAGTATTGGG	1619

QY	721	AGATCGATGGGCAAAACATCTGTAGATGACAGAACCGCTGGTCTTTTTCACAGACAG	780
DB	1620	AGATCGATGGGCAAAACATCTGTAGATGACAGAACCGCTGGTCTTTTTCACAGACAG	1679
QY	781	TTCTGACAGTGGAGAGCTCTGCACCTTTCTCTGAGGAACTTCTGGTGTGGCTACAGCT	840
DB	1680	TTCTGACAGTGGAGAGCTCTGCACCTTTCTCTGAGGAACTTCTGGTGTGGCTACAGCT	1739
QY	841	GAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCA	900
DB	1740	GAAAGATGATGAATTAAGCCGCGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCA	1799
QY	901	GCAGAACGATGTACATAGGCGCTTCAAGAGGAATTTGAAAACCTTAAGAACCTTGTATCAT	960
DB	1800	GCAGAACGATGTACATAGGCGCTTCAAGAGGAATTTGAAAACCTTAAGAACCTTGTATCAT	1859
QY	961	GAGTACTCTTTGAGACGTGACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA	1020
DB	1860	GAGTACTCTTTGAGACGTGACGAATATTTCTGACAGAGCAGCCCTTTGGAAGGACTAGAGAA	1919
QY	1021	ACTCTACAGGAGCCCGCAGAGAGCTGCTCTCTGAGGAGAGAGCCGAGAACTGTCTCGGCT	1080
DB	1920	ACTCTACAGGAGCCCGCAGAGAGCTGCTCTCTGAGGAGAGAGCCGAGAACTGTCTCGGCT	1979
QY	1081	TCTACGAAACAGCGCTGAGGAGGTCAATCTGAGTGGGAAAATTTGAACTGCACTCCGC	1140
DB	1980	TCTACGAAACAGCGCTGAGGAGGTCAATCTGAGTGGGAAAATTTGAACTGCACTCCGC	2039
QY	1141	TGACTGCGCAGAGAAAATAGATGAGACCCCTTTGAAAGACTTCCAGGAACTTCAAGAGGCCAC	1200
DB	2040	TGACTGCGCAGAGAAAATAGATGAGACCCCTTTGAAAGACTTCCAGGAACTTCAAGAGGCCAC	2099
QY	1201	GGATGAGCTGGAACCTCAAGCTGGCCCAAGTGAAGTCAAGGGATCTTGGCAGGCCGT	1260
DB	2100	GGATGAGCTGGAACCTCAAGCTGGCCCAAGTGAAGTCAAGGGATCTTGGCAGGCCGT	2159
QY	1261	GGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGAGAAAGTCAAGSCACTTGGAGG	1320
DB	2160	GGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCGAGAAAGTCAAGSCACTTGGAGG	2219
QY	1321	AGAAATTTGCGCTCTGAAAGAGAAAGTGAAGCAAGTCAATGACTCTTGGCTGGCAGCTTAC	1380
DB	2220	AGAAATTTGCGCTCTGAAAGAGAAAGTGAAGCAAGTCAATGACTCTTGGCTGGCAGCTTAC	2279
QY	1381	CACCTTTGGCGATTCAGCTCTCACCGTATAACCTCAGCACTCTCGAAAGACCTGAACACCAG	1440
DB	2280	CACCTTTGGCGATTCAGCTCTCACCGTATAACCTCAGCACTCTCGAAAGACCTGAACACCAG	2339
QY	1441	ATGGAAGCTTTCTGAGGTGGCCCTCGAGGACCGAGTCAAGTCAAGTCAAGTCAAGTCAAGT	1500
DB	2340	ATGGAAGCTTTCTGAGGTGGCCCTCGAGGACCGAGTCAAGTCAAGTCAAGTCAAGTCAAGT	2399
QY	1501	GGACTTTGGTCCAGCTCTCAGCACTTTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAG	1560
DB	2400	GGACTTTGGTCCAGCTCTCAGCACTTTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAG	2459
QY	1561	AGCCATCTCGCCAAACAAAGTGGCCCTTACTATATCAACCAAGACTCAAACTTGGCTG	1620
DB	2460	AGCCATCTCGCCAAACAAAGTGGCCCTTACTATATCAACCAAGACTCAAACTTGGCTG	2519
QY	1621	GGACCATCTCCAAATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATATGTACATT	1680
DB	2520	GGACCATCTCCAAATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATATGTACATT	2579
QY	1681	CTCAGCTTATAGGACTGCAATGAACTCCGAAAGACTCGAAGGCCCTTTGCTTGGATCT	1740
DB	2580	CTCAGCTTATAGGACTGCAATGAACTCCGAAAGACTCGAAGGCCCTTTGCTTGGATCT	2639
QY	1741	CTTGAGCTGTGAGCTGATGTGCTTGGACACCACTCAAGCAAAATGACCA	1800
DB	2640	CTTGAGCTGTGAGCTGATGTGCTTGGACACCACTCAAGCAAAATGACCA	2699

1801 GCCATGGATATCTCGAGATTATTATTTGTTGACCACTATTATGACCGCTGGAGCA 1860
2700 GCCATGGATATCTCGAGATTATTATTTGTTGACCACTATTATGACCGCTGGAGCA 2759
1861 AGAGCAACAATTGGTGTACAGTCCCTCTCTGCGTGGATATGTGTGAATCGCTGCT 1920
2760 AGAGCAACAATTGGTGTACAGTCCCTCTCTGCGTGGATATGTGTGAATCGCTGCT 2819
1921 GAATGTTTATGATACGGGACGAACAGGAGGATCCGTGCTCTCTTTTAAACTGGCAT 1980
2820 GAATGTTTATGATACGGGACGAACAGGAGGATCCGTGCTCTTTTAAACTGGCAT 2879
1981 CATTCCTCTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
2880 CATTCCTCTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGC 2939
2041 AAGTTCAACAGGATTTGTGACAGCGGAGGATCCGTGCTCTCTCTCTCATGATTTCTATCCA 2100
2940 AAGTTCAACAGGATTTGTGACAGCGGAGGATCCGTGCTCTCTCTCTCATGATTTCTATCCA 2999
2101 A 2101
3000 A 3000

RESULT 2
AAD37259
ID AAD37259 standard; DNA; 4476 BP.
XX
AC AAD37259;
XT 21-AUG-2002 (first entry)
DE Adeno-associated virus vector plasmid, AAV-MCK-3510.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX WO200183695-A2.
XX 08-NOV-2001.
XX 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
XX Xiao X;
XX WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
XX Example 1; Page 63-65; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV

CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 other;
Query Match 100.0%; Score 2101; DB 24; Length 4476;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGCTATGCTTACACACAGCGTCTTATGTACACACCTCTGACCTACACGAGCCCAATT 60
DB 1656 GAGCTATGCTTACACACAGCGTCTTATGTACACCTCTGACCTACACGAGCCCAATT 1715
QY 61 TCCTTCACAGCATTTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTTCATGTAGAGAG 120
DB 1716 TCCTTCACAGCATTTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTTCATGTAGAGAG 1775
QY 121 TGAAGTAAACCTGACCGTTTATCAACAGCTTTAGAAAGAGTATTATCGTGGCTTCCTTC 180
DB 1776 TGAAGTAAACCTGACCGTTTATCAACAGCTTTAGAAAGAGTATTATCGTGGCTTCCTTC 1835
QY 181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA 240
DB 1836 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA 1895
QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCGGGTTGG 300
DB 1896 CCAGTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCGGGTTGG 1955
QY 301 TAATATTTCTACATTTGGGAAGTAAAGCTGATTGGAACAGAGGAAATTTATCAGAAGATGAAGA 360
DB 1956 TAATATTTCTACATTTGGGAAGTAAAGCTGATTGGAACAGAGGAAATTTATCAGAAGATGAAGA 2015
QY 361 AACTGAAGTACAGAGCAGATGAATTCCTTAATTCAGATGGGAATCCCTCAGGTAGC 420
DB 2016 AACTGAAGTACAGAGCAGATGAATTCCTTAATTCAGATGGGAATCCCTCAGGTAGC 2075
QY 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
DB 2076 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 2135
QY 481 GAAAGAGTTGAATGATCGTGGCTTACAAAGAAACAGAGAAACAAAGGAAATGAGAGGAAGA 540
DB 2136 GAAAGAGTTGAATGATCGTGGCTTACAAAGAAACAGAGAAACAAAGGAAATGAGAGGAAGA 2195
QY 541 GCCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAGGTGCTTCA 600
DB 2196 GCCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAGGTGCTTCA 2255
QY 601 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 660
DB 2256 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCTCACTCACATGGTGGTGTAGT 2315
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAGAACAACTTAAGTATTGGG 720
DB 2316 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAGAACAACTTAAGTATTGGG 2375
QY 721 AGATCGATGGGCAACATCTGTAGATGGAAGAGACCGCTGGGTCTCTTTTACAAGACAG 780
DB 2376 AGATCGATGGGCAACATCTGTAGATGGAAGAGACCGCTGGGTCTCTTTTACAAGACAG 2435
QY 781 TTCTGACCAAGTGGAGCGTCTGCACTTTCTCTGACGAACTTCTGTGTGGCTTACAGCT 840
DB 2436 TTCTGACCAAGTGGAGCGTCTGCACTTTCTCTGACGAACTTCTGTGTGGCTTACAGCT 2495
QY 841 GAAAGATCATGAATTAAGCCGGCAGGCACTTATGGAGGCACTTTCCAGCAGTTTCAGAA 900
DB 2496 GAAAGATCATGAATTAAGCCGGCAGGCACTTATGGAGGCACTTTCCAGCAGTTTCAGAA 2555
QY 901 GCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTAAGAACTTAAGAACTGTGAATCAT 960
DB 2556 GCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTAAGAACTTAAGAACTGTGAATCAT 2615

2y 181 TGTGAGGACACATTCGAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
2b 1279 TGTGAGGACACATTCGAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1338
2y 241 CCAAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 300
2b 1339 CCAAGTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 1398
2y 301 TAATATTTCTACAAATTTGGGAAGTAAGTGTGATTTGGAAACAGGAAATTTATCAGAAAGATGAAGA 360
2b 1399 TAATATTTCTACAAATTTGGGAAGTAAGTGTGATTTGGAAACAGGAAATTTATCAGAAAGATGAAGA 1458
2y 361 AACTGAAGTACAGAGACAGATGATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 420
2b 1459 AACTGAAGTACAGAGACAGATGATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 1518
2y 421 TAGCATTTGGAATAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAAACT 480
2b 1519 TAGCATTTGGAATAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAAACT 1578
2y 481 GAAAGAGTTGAATGAATGGCTAAACAAAAACAGAGAAAGAAACAGGAAATGGAGGAAGA 540
2b 1579 GAAAGAGTTTCTTGGCTGCTTACAGAGCTGAAACAACTGCGCAATGTCTCAGAGATGC 1638
2y 541 GCCTCTTGGACCTGATCTTCAGACCTTAAACGCCCAAGTACACACATCAAGGTGCTTCA 600
2b 1639 TACCCGTAAAGAAAGGCTCTTAGAGACTTCAAGAGGATTAAGAGCTGATGAACAAATG 1698
2y 601 AGAAGATCTAGAAACAGAACTCAGGGTCAATTTCTCACTACATGGTGGTGTAGT 660
2b 1699 GCAAGACCTTCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACCTGGATGAAA 1758
2y 661 TGAATG-----ATCTAGTGGAGATCAGCACTGCTGCTGCTTTGGAGAACAACT 708
2b 1759 CAGCCAAAAAATCCTGAGATCCCTGGAAGGTTCCGATGATGCGATGCTCTTACAAAGAG 1818
2y 709 TAAGG---TATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGT 765
2b 1819 TTTGGATAACATGAATCTCAAGTGGAGTGAATCTCGAAGAAAGTCTCTCAACATTAGGTC 1878
2y 766 TCATTTACAGACAGTTCTGACAGTGGAGCGTCTGCACTTTCTCTGACAGAACTTCT 825
2b 1879 CCATTTGGAAGCCAGTTCTGACAGTGGAGCGTCTGCACTTTCTCTGACAGAACTTCT 1938
2y 826 GGTGTGCTACAGCTGAAGATCATGAATTAAGCGGAGGACCTATTGGAGGGGACTT 885
2b 1939 GGTGTGCTACAGCTGAAGATCATGAATTAAGCGGAGGACCTATTGGAGGGGACTT 1998
2y 886 TCCAGCAGTTTCAAGACAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTAA 945
2b 1999 TCCAGCAGTTTCAAGACAGAACGATGTACATAGGGCCCTTCAAGAGGGAATTGAAACTAA 2058
2y 946 AGAACCTGTAATCATGATGATCTCTGAGACTGTACGAATATTCTGACAGAGCAGCTTT 1005
2b 2059 AGAACCTGTAATCATGATGATCTCTGAGACTGTACGAATATTCTGACAGAGCAGCTTT 2118
2y 1006 GGAAGGACTAGAGAACTCTTACAGAGGCGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCA 1065
2b 2119 GGAAGGACTAGAGAACTCTTACAGAGGCGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCA 2178
2y 1066 GAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAATTT 1125
2b 2179 GAATGTCACTCGGCTTCTAGAAAGCAGGCTGAGGAGGTCAATACTAGTGGGAAATTT 2238
2y 1126 GAACCTCACTCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1185
2b 2239 GAACCTCACTCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2298
2y 1186 ACTTCAAGAGCCACGATGAGTGGACCTCAAGCTGCGCAAGAGCTGAGTGTATCAAGGG 1245
2b 2299 ACTTCAAGAGCCACGATGAGTGGACCTCAAGCTGCGCAAGAGCTGAGTGTATCAAGGG 2358
2y 1246 ATCTGCGAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAGATCACCTCGAGAACT 1305

2b 2359 ATCTGCGAGCCCGTGGCGATCTCTCTCATTTGACTCTCTCCAGATCACCTCGAGAACT 2418
2y 1306 CAAGGCATCTTCGAGGAGAAATTCGGCTCTGAAAGAGAAAGCTGAGCCACGCTCAATGACCT 1365
2b 2419 CAAGGCATCTTCGAGGAGAAATTCGGCTCTGAAAGAGAAAGCTGAGCCACGCTCAATGACCT 2478
2y 1366 TGTCTGCGACCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAATCACTCAGCATCTGGA 1425
2b 2479 TGTCTGCGACCTTACCACTTTGGGCAATTCAGCTCTCACCGTATAAATCACTCAGCATCTGGA 2538
2y 1426 AGACTCAACACACAGATGGAAGCTTTGCAAGTGGCGTCCGAGACCGAGTCAAGGACGCT 1485
2b 2539 AGACTCAACACACAGATGGAAGCTTTGCAAGTGGCGTCCGAGACCGAGTCAAGGACGCT 2598
2y 1486 GCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCATCTTCTTTCCACGCTGTGTCCA 1545
2b 2599 GCATGAAGCCCAACAGGACCTTTGGTCCAGCATCTCAGCATCTTCTTTCCACGCTGTGTCCA 2658
2y 1546 GGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGAC 1605
2b 2659 GGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGAC 2718
2y 1606 TCAAAACAACTTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCT 1665
2b 2719 TCAAAACAACTTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAAGTCTTTAGCTGACCT 2778
2y 1666 GAATAATGTGATCTCAGCTTATAGGACTGCATGAATACTCGAAGACTCGAGAGGC 1725
2b 2779 GAATAATGTGATCTCAGCTTATAGGACTGCATGAATACTCGAAGACTCGAGAGGC 2838
2y 1726 CCTTTGCTTGGATCTCTTGGAGCTGTGATGATGCTTGGACCAAGCAACAACT 1785
2b 2839 CCTTTGCTTGGATCTCTTGGAGCTGTGATGATGCTTGGACCAAGCAACAACT 2898
2y 1786 CAAGCAAAATGACAGCCCATGATATCTGCGAGATTTAATTTGTTGACCACTATTTA 1845
2b 2899 CAAGCAAAATGACAGCCCATGATATCTGCGAGATTTAATTTGTTGACCACTATTTA 2958
2y 1846 TGACCGCTCGAGCAAGCAGACCAACAACTTTGGTCAACGCTCTCTGCGTGGATATGTG 1905
2b 2959 TGACCGCTCGAGCAAGCAGACCAACAACTTTGGTCAACGCTCTCTGCGTGGATATGTG 3018
2y 1906 TCTGAATGCTGCTGAATTTATGATACGGACGAAACAGGAGGATCCGTGCTGTGTC 1965
2b 3019 TCTGAATGCTGCTGAATTTATGATACGGACGAAACAGGAGGATCCGTGCTGTGTC 3078
2y 1966 TTTTAAACTGGCATCATTTCCCTGCTAAAGCACATTTTGGAGACAAAGTACAGATACCT 2025
2b 3079 TTTTAAACTGGCATCATTTCCCTGCTAAAGCACATTTTGGAGACAAAGTACAGATACCT 3138
2y 2026 TTTCAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGAGGCTGGGCTCTCTTCT 2085
2b 3139 TTTCAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGAGGCTGGGCTCTCTTCT 3198
2y 2086 GCATGATTTCTATCCAA 2101
2b 3199 GCATGATTTCTATCCAA 3214

RESULT 4

AAD37238 standard; DNA; 3531 BP.
ID AAD37238

XX AAD37238;

AC AAD37238;

DT 21-AUG-2002 (first entry)

DE Human dystrophin minigene delta3531.

Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.

```
XX OS Homo sapiens.
XX PN WO200183695-A2.
XX PD 08-NOV-2001.
XX PF 27-APR-2001; 2001WO-US13677.
XX PR 28-APR-2000; 2000US-20077P.
XX PA (XIAO/) XIAO X.
XX PI Xiao X;
XX DR WPI; 2002-049342/06.
XX PT New dystrophin minigene for treating Duchenne or Becker muscular
XX PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX PT dystrophin gene.
XX PS Example 1; Page 50-51; 71pp; English.
XX CC The present invention relates to an isolated nucleotide sequence encoding
XX CC a dystrophin minigene. The minigene comprises N-terminal or modified
XX CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX CC invention also relates to a recombinant adeno-associated virus (AAV)
XX CC comprising dystrophin minigene operably linked to an expression control
XX CC element. The dystrophin minigene in operable linkage with an expression
XX CC control element, in a recombinant adeno-associated virus or retrovirus is
XX CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX CC dystrophy (BMD) in a mammalian subject. The present sequence is human
XX CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
XX CC hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and
XX CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX SQ Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 other;
Query Match 83.3%; Score 1750.8; DB 24; Length 3531;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;
QY 1 GAGCTATGCTACACAGAGCTCTTATCTCACCCTCTGACCTTACACGAGCCCAT 60
DB 900 GAGCTATGCTACACAGAGCTCTTATCTCACCCTCTGACCTTACACGAGCCCAT 959
QY 61 TCCTTCACAGCATTTGGAAGCTCTCTCAAGACAAGTCATTTGGCAGTTTCATTGGAGAG 120
DB 960 TCCTTCACAGCATTTGGAAGCTCTCTCAAGACAAGTCATTTGGCAGTTTCATTGGAGAG 1019
QY 121 TGAAGTAACTGGACCGTTTCAACACAGCTTTAGAAAGATTTATTCGGCTTCTTTC 180
DB 1020 TGAAGTAACTGGACCGTTTCAACACAGCTTTAGAAAGATTTATTCGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTGCAAGCAAGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
DB 1080 TGCTGAGGACACATTGCAAGCAAGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCCGGGTGG 300
DB 1140 CCAGTTTTCATCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCCGGGTGG 1199
QY 301 TAATATTCTCAATTTGGAAGTAACTGATTGGAAACAGGAAATTTATCAGAAGATCAAGA 360
DB 1200 TAAATTCTCAATTTGGAAGTAACTGATTGGAAACAGGAAATTTATCAGAAGATCAAGA 1259
QY 361 AACTCAAGTACAGAGCAGATGAAATCTCTTAAATTTCAAGATGGGAATGCTCAGGTAGC 420
DB 1260 AACTCAAGTACAGAGCAGATGAAATCTCTTAAATTTCAAGATGGGAATGCTCAGGTAGC 1319
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGATT-----TTAATGGATCTCCAGATCA 474
```

```
1320 TAGCATGGAAAAACAAGCAATTTACATAGAACTATAGATTACTGCAACAGTTCCCCCT 1379
475 GAACTGGAAGAGTTGAATGACTGCTAAACAAAACAGAAAGAAACAAGGAAATCGA 534
1380 GGACCTGGAAAGTTTCTTGCTGGCTTACAGAGCTGAAACAACTGCAATTCCTACA 1439
535 GGAAGAGCTCTTGACCTGATCTTGAAGACCTAAACGCCAAGTACAAACACATAGGT 594
1440 GGATGCTACCCGTAAGGAAAGGCTCTTAGAAGACTTCAAGGGAGTAAAGAGCTGATGAA 1499
595 GCTTCAAGNAGATCTAGAACAGAACNAGTCAGGTCATTTCTCTCACTCACATGGTGT 654
1500 ACAATGGCAAGACCTCCAAAGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGA 1559
655 GGTAG-----TTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGA 702
1560 TGAACACAGCCAAAAATCTCGAGATCCTGGAAGTTCCGATGATGAGTCTCTTACA 1619
703 ACACTTAAGG---TATTGGGAGATCGATGGCAACATCTGTAGATGGACAGAGACCG 759
1620 AAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACAT 1679
760 CTGGGTTCTTTTACAAAGACAGTTTCTGACCAGTGAAGCGTCTGCACTTTCTCTCAGGA 819
1680 TAGTCTCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCACTTTCTCTCAGGA 1739
820 ACTTCTGGTGGCTACACTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 879
1740 ACTTCTGGTGGCTACACTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 1799
880 CGACTTTCAGCAGTTTCAGAAGCAGAACCATGTACATAGGGCTTCAAGAGGGAATTGAA 939
1800 CGACTTTCAGCAGTTTCAGAAGCAGAACCATGTACATAGGGCTTCAAGAGGGAATTGAA 1859
940 AACTAAAGAACCTGTAAATCATAGATGATCTTGGAGCTGACGAATTTCTGACGAGCA 999
1860 AACTAAAGAACCTGTAAATCATAGATGATCTTGGAGCTGACGAATTTCTGACGAGCA 1919
1000 GCCTTTGGAAGGACTAGAGAACTCTTACAGAGGCCAGAGAGCTCCCTCCTGAGGAGAG 1059
1920 GCCTTTGGAAGGACTAGAGAACTCTTACAGAGGCCAGAGAGCTCCCTCCTGAGGAGAG 1979
1060 AGCCCAAGATGTCACTCGGCTTCTAGAAAGCGGCTGAGGAGTCAATCTGAGTGGGA 1119
1980 AGCCCAAGATGTCACTCGGCTTCTAGAAAGCGGCTGAGGAGTCAATCTGAGTGGGA 2039
1120 ARAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAATAATAGATGAGACCTTGAAGACT 1179
2040 ARAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAATAATAGATGAGACCTTGAAGACT 2099
1180 CCAGGAATTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGTGTAT 1239
2100 CCAGGAATTTCAAGAGGCCACCGATGAGTGGACCTCAAGCTCGGCCAAGCTGAGTGTAT 2159
1240 CAAGGATCTTGGCAGCCCGTGGCGATCTCTCACTATTGACTCTCTCAAGATCACTCGA 1299
2160 CAAGGATCTTGGCAGCCCGTGGCGATCTCTCACTATTGACTCTCTCAAGATCACTCGA 2219
1300 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAA 1359
2220 GAAAGTCAAGGCACTTCGAGGAGAAATTTGGCCCTCTGAAAGAGAAACGTGAGCCACGTCAA 2279
1360 TGAOCTTGTCTGCCAGCTTACCATTGGGGATTCAGCTCTCCCGTATTAACCTCAGCAC 1419
2280 TGAOCTTGTCTGCCAGCTTACCATTGGGGATTCAGCTCTCCCGTATTAACCTCAGCAC 2339
1420 TCTGGAAGACCTGAAACCAAGATGGAAGCTTCTCAGGTGGCGCTCGAGGACCGAGTCAG 1479
2340 TCTGGAAGACCTGAAACCAAGATGGAAGCTTCTCAGGTGGCGCTCGAGGACCGAGTCAG 2399
1480 GCAGTGCATGAAGCCCAAGGAGCTTTGGTTCAGGATCTCTCAGCACTTTCTTCCAGCTC 1539
```


Db 2400 GCAGCTCATGAAAGCCACAGGACCTTTGGTCCAGCATCTCAGACCTCTTCTTCCACGTC 2459
Qy 1540 TGTCCAGGTCCTGGAGAGACCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 1599
Db 2460 TGTCCAGGTCCTGGAGAGACCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 2519
Qy 1600 CGAGACTCAAAACACTTGTCTGGAGACCATCCCAAAATGACAGAGCTCTACAGTCCTTTAGC 1659
Db 2520 CGAGACTCAAAACACTTGTCTGGAGACCATCCCAAAATGACAGAGCTCTACAGTCCTTTAGC 2579
Qy 1660 TGACCTGAATAATGTAGATCTCAGCTTATAGAGTCCCATGAACTCGGAAGCTGCA 1719
Db 2580 TGACCTGAATAATGTAGATCTCAGCTTATAGAGTCCCATGAACTCGGAAGCTGCA 2639
Qy 1720 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACACAGCA 1779
Db 2640 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACACAGCA 2699
Qy 1780 CAACCTCAAGCAAAATGACAGGCCCATCGATATCCTGTCAGATTAATTTGTTGACCCAC 1839
Db 2700 CAACCTCAAGCAAAATGACAGGCCCATCGATATCCTGTCAGATTAATTTGTTGACCCAC 2759
Qy 1840 TATTTATGACCGCTGGAGAGACCAACAATTTGTCACAGTCCCTCTCTGCGTGGCA 1899
Db 2760 TATTTATGACCGCTGGAGAGACCAACAATTTGTCACAGTCCCTCTCTGCGTGGCA 2819
Qy 1900 TATGTGTCAACTGGCTGCTGTAATTTATGATACGGGACGACAGGGAGGATCCGCTGT 1959
Db 2820 TATGTGTCAACTGGCTGCTGTAATTTATGATACGGGACGACAGGGAGGATCCGCTGT 2879
Qy 1960 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGACATTTGGAGACAAATACAG 2019
Db 2880 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGACATTTGGAGACAAATACAG 2939
Qy 2020 ATACCTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACGCGCAGGCTGGGCGCT 2079
Db 2940 ATACCTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACGCGCAGGCTGGGCGCT 2999
Qy 2080 CCTTCTGCATCAATCTATCCAA 2101
Db 3000 CCTTCTGCATCAATCTATCCAA 3021

RESULT 5

AAD37258
ID AAD37258 standard; DNA; 4498 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX

XX Example 1; Page 62-63; 7lpp; English.

XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid construct containing human dystrophin minigenes, a muscle
XX creatine kinase (MCK) promoter and a small polyA signal sequence.

XX Sequence 4498 BP; 1251 A; 1118 C; 1123 G; 1006 T; 0 other;

Query Match 83.3%; Score 1750.8; DB 24; Length 4498;

Best Local Similarity 90.2%; Pred. No. 0;

Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;

Qy 1 GAGCTATGCCCTACACACAGGCTGCTTATGTCCACCTCTGACCTTACAGGAGCCCATTT 60

Db 1657 GAGCTATGCCCTACACACAGGCTGCTTATGTCCACCTCTGACCTTACAGGAGCCCATTT 1716

Qy 51 TCCCTTACAGCATTTGGAAGCTCTGGAAGCAAGTTCATTTGGCAGTTTCATTGATGGAGAG 120

Db 1717 TCCCTTACAGCATTTGGAAGCTCTGGAAGCAAGTTCATTTGGCAGTTTCATTGATGGAGAG 1776

Qy 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTTGAAGAAGTATTATCGTGGCTTCTTTTC 180

Db 1777 TGAAGTAAACCTGGACCGTTATCAACAGCTTTTGAAGAAGTATTATCGTGGCTTCTTTTC 1836

Qy 181 TGCTGAGGACACATTTGCAAGCAGCAGGAGAGATTTCTAATGATGTGAAGTGGTGAAGA 240

Db 1837 TGCTGAGGACACATTTGCAAGCAGCAGGAGAGATTTCTAATGATGTGAAGTGGTGAAGA 1896

Qy 241 CCAGTTTTCATCTCATGAGGGGTACATGATGATGATTTGA CAGCCCATCAGGCCCGGTTGG 300

Db 1897 CCAGTTTTCATCTCATGAGGGGTACATGATGATGATTTGA CAGCCCATCAGGCCCGGTTGG 1956

Qy 301 TAAATTTCTCAATTTGGGAAGTAACTGATTTGGAACAGGAAAATTAATCAGAAGATGAAGA 360

Db 1957 TAAATTTCTCAATTTGGGAAGTAACTGATTTGGAACAGGAAAATTAATCAGAAGATGAAGA 2016

Qy 361 RACTCAAGTACAAAGACAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 420

Db 2017 RACTCAAGTACAAAGACAGATGAATCTCTTAAATTCAGATGGGAATGCCTCAGGGTAGC 2076

Qy 421 TAGCATGGAATAAACAAGCAATTTACATAGAGTT-----TTAATGGATCTCCAGAAATCA 474

Db 2077 TAGCATGGAATAAACAAGCAATTTACATAGAGTT-----TTAATGGATCTCCAGAAATCA 2136

Qy 475 GAAACTGGAAGAGTTGAATGACTGGCTTAACAAAACAGAGAAAGAAACAAAGAAATGGA 534

Db 2137 GAACTGGAAGAGTTGAATGACTGGCTTAACAAAACAGAGAAAGAAACAAAGAAATGGA 2196

Qy 535 GGAAGAGCCTCTTTGGACCTGTATCTTTGAAGACCTTAAACGCCAAGTACAAACATAAGGT 594

Db 2197 GGATGCTACCCGTAAGGAAAGGCTCTTAGAGACTCCAAAGGAGTAAAGAGCTGATGAA 2256

Qy 595 GCTTCAAGAGATCTTAGAACAAGAACAGTACAGGTCATTTCTCTCCTCATCATGCTGCT 654

Db 2257 ACAATGGCAAGACCTCCAAAGGTGAATTTGAAGCTCACAGATGTTTATCACAACCTGGA 2316

Qy 655 GGTAG-----TTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGA 702

Db 2400 GCAGCTCATGAAAGCCACAGGACCTTTGGTCCAGCATCTCAGACCTCTTCTTCCACGTC 2459

Qy 1540 TGTCCAGGTCCTGGAGAGACCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 1599

Db 2460 TGTCCAGGTCCTGGAGAGACCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 2519

Qy 1600 CGAGACTCAAAACACTTGTCTGGAGACCATCCCAAAATGACAGAGCTCTACAGTCCTTTAGC 1659

Db 2520 CGAGACTCAAAACACTTGTCTGGAGACCATCCCAAAATGACAGAGCTCTACAGTCCTTTAGC 2579

Qy 1660 TGACCTGAATAATGTAGATCTCAGCTTATAGAGTCCCATGAACTCGGAAGCTGCA 1719

Db 2580 TGACCTGAATAATGTAGATCTCAGCTTATAGAGTCCCATGAACTCGGAAGCTGCA 2639

Qy 1720 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACACAGCA 1779

Db 2640 GAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTGAGCTGCATGTGATGCCCTTGACACAGCA 2699

Qy 1780 CAACCTCAAGCAAAATGACAGGCCCATCGATATCCTGTCAGATTAATTTGTTGACCCAC 1839

Db 2700 CAACCTCAAGCAAAATGACAGGCCCATCGATATCCTGTCAGATTAATTTGTTGACCCAC 2759

Qy 1840 TATTTATGACCGCTGGAGAGACCAACAATTTGTCACAGTCCCTCTCTGCGTGGCA 1899

Db 2760 TATTTATGACCGCTGGAGAGACCAACAATTTGTCACAGTCCCTCTCTGCGTGGCA 2819

Qy 1900 TATGTGTCAACTGGCTGCTGTAATTTATGATACGGGACGACAGGGAGGATCCGCTGT 1959

Db 2820 TATGTGTCAACTGGCTGCTGTAATTTATGATACGGGACGACAGGGAGGATCCGCTGT 2879

Qy 1960 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGACATTTGGAGACAAATACAG 2019

Db 2880 CCTGTCTTTAAACTGGCATCATTTCCCTGTGTAAAGACATTTGGAGACAAATACAG 2939

Qy 2020 ATACCTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACGCGCAGGCTGGGCGCT 2079

Db 2940 ATACCTTTCAAGCAAGTGGCAAGTTCACAGAGATTTTGTGACGCGCAGGCTGGGCGCT 2999

Qy 2080 CCTTCTGCATCAATCTATCCAA 2101

Db 3000 CCTTCTGCATCAATCTATCCAA 3021

XX Adeno-associated virus vector plasmid, AAV-MCK-3531.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

XX Becker muscular dystrophy; ds.

XX Chimeric - Homo sapiens.

XX Chimeric - Unidentified.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

Db 2317 TGAACAACGCAAAATCTCGATCCCTGAGATCCCTGAGGTTCCGATGATGAGTCTCTTACA 2376
Qy 703 ACAACTTAAGG---TATTGGGAGATCGATGGGCAAAACATCTCTAGATGGACAGAAAGCCG 759
Db 2377 AAGACGTTTGGATTAACATGAATCTCAAGTGGAGTGAACCTTCGGAAAAAGTCTCTCAACAT 2436
Qy 760 CTGGGTTCTTTACAGACAGTCTGACACAGTGGAGCGTCTGACCTTTCTCTCCAGGA 819
Db 2437 TAGGTCCCAATTTGGAAGCCAGTTCTGACACAGTGGAGCGTCTGACCTTTCTGACAGGA 2496
Qy 820 ACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 879
Db 2497 ACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGG 2556
Qy 880 CGACTTTCAGGAGTTTCAGACAGAACAGTGTACATAGGGCTTCAGAGGGAATTGAA 939
Db 2557 CGACTTTCAGGAGTTTCAGACAGAACAGTGTACATAGGGCTTCAGAGGGAATTGAA 2616
Qy 940 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 999
Db 2617 AACTAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCA 2676
Qy 1000 GCCTTTGGAAGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTGAGGAGAG 1059
Db 2677 GCCTTTGGAAGACTAGAGAACTCTACAGAGCCAGAGAGCTGCTCTGAGGAGAG 2736
Qy 1060 AGCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATPACTGAGTGGGA 1119
Db 2737 AGCCAGATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTCAATPACTGAGTGGGA 2796
Qy 1120 AAAATTGAACCTGCACTCCGCTGACTGGGAGAGAAAAATAGATGAGACCTTTGAAGACT 1179
Db 2797 AAAATTGAACCTGCACTCCGCTGACTGGGAGAGAAAAATAGATGAGACCTTTGAAGACT 2856
Qy 1180 CCAGGAATCTCAAGAGCCAGGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 1239
Db 2857 CCAGGAATCTCAAGAGCCAGGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGAT 2916
Qy 1240 CAAGGGATCTTGGAGCCGCTGGGCGATCTCTCATTTGACTCTCTCAAGATCACTCGA 1299
Db 2917 CAAGGGATCTTGGAGCCGCTGGGCGATCTCTCATTTGACTCTCTCAAGATCACTCGA 2976
Qy 1300 GAAAGTCAAGCACTTCGAGAGAAATTTGGCGCTCTGAAAGAGAACTGAGCCACCTCAA 1359
Db 2977 GAAAGTCAAGCACTTCGAGAGAAATTTGGCGCTCTGAAAGAGAACTGAGCCACCTCAA 3036
Qy 1360 TGACCTTGTCTCGCCAGCTTACCATTGTTGGCATTTGAGCTCTCACCGTATAAAGCTCAGCAC 1419
Db 3037 TGACCTTGTCTCGCCAGCTTACCATTGTTGGCATTTGAGCTCTCACCGTATAAAGCTCAGCAC 3096
Qy 1420 TCTGGAAGACTGAAACACAGATGGAAGCTCTGAGGTGGCGCTCGAGGACCGAGTCAAG 1479
Db 3097 TCTGGAAGACTGAAACACAGATGGAAGCTCTGAGGTGGCGCTCGAGGACCGAGTCAAG 3156
Qy 1480 CGAGCTGATGAAGCCACAGGAGTCTTGGTCCAGCATCTCAGACCTTTCTTTCCACGTC 1539
Db 3157 CGAGCTGATGAAGCCACAGGAGTCTTGGTCCAGCATCTCAGACCTTTCTTTCCACGTC 3216
Qy 1540 TGTCAGGCTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 1599
Db 3217 TGTCAGGCTCCCTGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACCA 3276
Qy 1600 CGAGACTCAAAACATTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 1659
Db 3277 CGAGACTCAAAACATTTGCTGGGACCATCCCAAAATGACAGAGCTTACCAGTCTTTAGC 3336
Qy 1660 TGACCTGAATTAATGTGAGATTTCTAGCTTATAGACTGCGATGAACTCCGAGACTGCA 1719
Db 3337 TGACCTGAATTAATGTGAGATTTCTAGCTTATAGACTGCGATGAACTCCGAGACTGCA 3396
Qy 1720 GAAGGCCCTTTGCTGGATCTCTTGGAGCTCTCAGCTGCAATGTGATGCTTTGAGACAGCA 1779
Db 3397 GAAGGCCCTTTGCTGGATCTCTTGGAGCTCTCAGCTGCAATGTGATGCTTTGAGACAGCA 3456

Qy 1780 CAACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTAATTTGTTGACCAC 1839
Db 3457 CAACCTCAAGCAAAATGACAGCCCATGATATCTCTGAGATTAATTAATTTGTTGACCAC 3516
Qy 1840 TATTATGACCGCTCGAGCAGAGACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 1899
Db 3517 TATTATGACCGCTCGAGCAGAGACAAATTTGGTCAAGCTCCCTCTCTGCGTGA 3576
Qy 1900 TATGTGTCTGAACCTGCTGCTGAATCTTTATGATACCGACCAACAGGAGGATCCGTCT 1959
Db 3577 TATGTGTCTGAACCTGCTGCTGAATCTTTATGATACCGACCAACAGGAGGATCCGTCT 3636
Qy 1960 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAGTACAG 2019
Db 3637 CCTGTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAGTACAG 3696
Qy 2020 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGCAGCGCTGGGCT 2079
Db 3697 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACCGCAGCGCTGGGCT 3756
Qy 2080 CCTTCTGCATGATTTCTATCCAA 2101
Db 3757 CCTTCTGCATGATTTCTATCCAA 3778

RESULT 6

AAD37237
ID AAD37237 standard; DNA; 3858 BP.
XX
AC AAD37237;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
PS WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
Example 1; Page 48-49; 71pp; English.
XX
The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified
N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is human

CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;
Query Match 83.0%; Score 1743; DB 24; Length 3858;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 GAGTATGCTACACACAGGCTGCTTATGTACACACCTCTGACCTACACGGAGCCCAT 60
DB GAGTATGCTTACACACAGGCTGCTTATGTACACACCTCTGACCTACACGGAGCCCAT 599
QY 61 TCCCTTACAGCATTTGGAAGCTCTGAAGACAAAGTCAATTTGGCAGTTTCAATGATGAGAG 120
DB TCCCTTACAGCATTTGGAAGCTCTGAAGACAAAGTCAATTTGGCAGTTTCAATGATGAGAG 1019
QY 121 TGAAGTAAACCTGGAACCGTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTTCTTTC 180
DB TGAAGTAAACCTGGAACCGTTATCAAAACAGCTTTAGAAAGATATATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACATTTGCAAGCACAGGAGAGATTTCTAATCATGTGGAAGTGTGAAAGA 240
DB TGCTGAGGACATTTGCAAGCACAGGAGAGATTTCTAATCATGTGGAAGTGTGAAAGA 1139
QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGGTTGG 300
DB CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGGTTGG 1199
QY 301 TAATATTCTCAATTTGGAGTAGTACCTCTTAAATTTCAAGATGGGAATTCAGAAATGAAGA 360
DB TAATATTCTCAATTTGGAGTAGTACCTCTTAAATTTCAAGATGGGAATTCAGAAATGAAGA 1259
QY 361 AACTCAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATTCAGAAATGAAGA 420
DB AACTCAAGTACAAGAGCAGATGAATCTCTTAAATTTCAAGATGGGAATTCAGAAATGAAGA 1319
QY 421 TAGCATGGAAGAACCAAGCAATTTACATGAGTTTAAATGATCTCCAGAACTCAGAACT 480
DB TAGCATGGAAGAACCAAGCAATTTACATGAGTTTAAATGATCTCCAGAACTCAGAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAAGCAAGGAAAATGGAGGAAGA 540
DB GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAAGCAAGGAAAATGGAGGAAGA 1439
QY 541 GCCTTTGGAACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACATTAAGGTGCTTCA 600
DB GCCTTTGGAACCTGATCTTGAAGACCTTAAACCGCCAAAGTACAAACATTAAGGTGCTTCA 1499
QY 601 AGAAGATCTAGAACAGAACCAAGTCAAGTCAATTTCTCACTCAGATGGTGGTGGTAGT 660
DB AGAAGATCTAGAACAGAACCAAGTCAAGTCAATTTCTCACTCAGATGGTGGTGGTAGT 1559
QY 661 TGATGAATCTAGTGAGATACGCAACTGCTGTTTGGAGAACCAACTTAAGGTATTTGGG 720
DB TGATGAATCTAGTGAGATACGCAACTGCTGTTTGGAGAACCAACTTAAGGTATTTGGG 1619
QY 721 AGATCGATGGGCAACATCTGTAGTACAGAGACCGCTGGGTTCTTTTACAGAG --- 777
DB AGATCGATGGGCAACATCTGTAGTACAGAGACCGCTGGGTTCTTTTACAGAG --- 777
QY 778 --- 777
DB 1680 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAGTTTCTTGGCTTTACAGA 1739
QY 778 --- 777
DB 1740 AGCTGAAACAACTGCCAATGTCTACAGGATGCTACCGTAAAGAAAGGCTCTTGAAGA 1799
QY 778 --- 777
DB 1800 CTCGAAGGAGTAAAGAGCTGATGAAACAAATGGCAAGACCTCCCAAGGTGAATTTGAAGC 1859

QY 778 --- 777
DB 1860 TCACACAGATGTTTATCACAACTTGATGAACACAGCCAAAATACTCTGAGATCCCTGGA 1919
QY 778 --- 777
DB 1920 AGGTTCCGATGATGCAGTCTGTGTACAAAGACGTTTGGATTAACATGAATCTCAAGTGAG 1979
QY 778 --- CAGTTCTGACCACTG 792
DB 1980 TGAACCTCGGAAAAGTCTCTCAACATTAGTCCCATTTGGGAAGCCAGTTCTGACCACTG 2039
QY 793 GAAGCGTCTGCACTTTCTCTGCGAGAACTTCTGGTGGCTACAGCTGAAGATGATGA 852
DB GAAGCGTCTGCACTTTCTCTGCGAGAACTTCTGGTGGCTACAGCTGAAGATGATGA 2099
QY 853 ATTAAGCCGGCAGGACCACTATTGGAGGGCACTTCCAGCAGTTTCCAGAGCAGTTCAGAAAGCAGACGATGT 912
DB ATTAAGCCGGCAGGACCACTATTGGAGGGCACTTCCAGCAGTTTCCAGCAGTTCAGAAAGCAGACGATGT 2159
QY 913 ACATAGGCGCTTCAAGAGGGAATTAAGAACTTAAAGAACTTGAATCATGATCATCTTTGA 972
DB ACATAGGCGCTTCAAGAGGGAATTAAGAACTTAAAGAACTTGAATCATGATCATCTTTGA 2219
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACCAAGGA 1032
DB GACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACCAAGGA 2279
QY 1033 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGTCTACTCGGCTTTTACGAAAGCA 1092
DB GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGTCTACTCGGCTTTTACGAAAGCA 2339
QY 1093 GCCTGAGAGGTCATCTGAGTGGGAAAATTTGAACCTGCACCTCCGCTGACTCGCAGAG 1152
DB GCCTGAGAGGTCATCTGAGTGGGAAAATTTGAACCTGCACCTCCGCTGACTCGCAGAG 2399
QY 1153 AAAAATAGATGAGACCTTTGAAAGACTCCAGAGAACTTCAAGAGGCCACGAGTGAAGTGA 1212
DB AAAAATAGATGAGACCTTTGAAAGACTCCAGAGAACTTCAAGAGGCCACGAGTGAAGTGA 2459
QY 1213 CTTCAAGCTGGCCAAAGCTGAGGTGATCAAGGATCTTGGCAGCCGCTGGGCGATCTCTCT 1272
DB CTTCAAGCTGGCCAAAGCTGAGGTGATCAAGGATCTTGGCAGCCGCTGGGCGATCTCTCT 2519
QY 1273 CATTCACCTCTCTCCAAAGTCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGGCC 1332
DB CATTCACCTCTCTCCAAAGTCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGGCC 2579
QY 1333 TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTCGCAGCTTACCACTTTGGGCGAT 1392
DB TCTGAAAGAGAACGTGAGCCACGTCATGACCTTGTCTCGCAGCTTACCACTTTGGGCGAT 2639
QY 1393 TCAGCTCTCCGCTATACCTCAGCAGCTCTCGAGAGCTGAAACCAACCATGAGAGCTTCT 1452
DB TCAGCTCTCCGCTATACCTCAGCAGCTCTCGAGAGCTGAAACCAACCATGAGAGCTTCT 2699
QY 1453 GCAGGTGGCCGCTCGAGGAGCCGAGTCAGGAGCTGATGAAGCCCAACAGGAGCTTTGGTCC 1512
DB GCAGGTGGCCGCTCGAGGAGCCGAGTCAGGAGCTGATGAAGCCCAACAGGAGCTTTGGTCC 2759
QY 1513 AGCATCTCAGCAGCTTTCTTTTCCAGCTGTCTCGAGGTCCTTGGGAGAGAGCCATCTCGCC 1572
DB AGCATCTCAGCAGCTTTCTTTTCCAGCTGTCTCGAGGTCCTTGGGAGAGAGCCATCTCGCC 2819
QY 1573 AAAACAAAGTGCCTACTATATCAACACAGAGACTCAAAACAACTTGTCTGGGACCATCCCAA 1632
DB AAAACAAAGTGCCTACTATATCAACACAGAGACTCAAAACAACTTGTCTGGGACCATCCCAA 2879
QY 1633 AATGACAGAGCTCTACCAAGTCTTTTAGCTGACCTGAAATTAATGTAGATTTCTCAGCTTATAG 1692
DB AATGACAGAGCTCTACCAAGTCTTTTAGCTGACCTGAAATTAATGTAGATTTCTCAGCTTATAG 2939

1693 GACTGCCATGAACCTCCGAAAGACTGAGAGGCCCTTTGCTTGGATCTCTTGAGCCTGTC 1752
1753 AGCTGCATGTGATGCTTGGACGAGCAGCAACCTCAAGCAAAATGACAGCCCATGGATAT 1812
3000 AGCTGCATGTGATGCTTGGACGAGCAGCAACCTCAAGCAAAATGACAGCCCATGGATAT 3059
1813 CTGCGAGATTAATTAATTTGACCACTATTATAGACCGCTGGAGCAAGACCAACAA 1872
3060 CTGCGAGATTAATTAATTTGACCACTATTATAGACCGCTGGAGCAAGACCAACAA 3119
1873 TTTGGTCAACGCTCCCTCTGCTGATATGCTCTGAACCTGCTGCTGCTGAATGTTATGA 1932
3120 TTTGGTCAACGCTCCCTCTGCTGATATGCTCTGAACCTGCTGCTGCTGAATGTTATGA 3179
1933 TACGGACGAAACAGGAGGATCGCTGCTGCTTTTAAACCTGGATCATTTCCCTGTG 1992
3180 TACGGACGAAACAGGAGGATCGCTGCTGCTTTTAAACCTGGATCATTTCCCTGTG 3239
1993 TAAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGG 2052
3240 TAAAGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGG 3299
2053 ATTTTGTGACACGCGAGGCTGGGCTCCTCTGCTGATGATTTCTATCCAA 2101
3300 ATTTTGTGACACGCGAGGCTGGGCTCCTCTGCTGATGATTTCTATCCAA 3348

RESULT 7
AAD37257
ID AAD37257 standard; DNA; 4825 BP.
XX
AC AAD37257;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
FI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 61-62; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control

CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a muscle
CC creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
SQ Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 other;
Query Match 83.0%; Score 1743; DB 24; Length 4825;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 GAGCTATGCTCACACAGGCTGCTTATGTCAACCTCTGACCCCTACACGAGCCCAT 60
Db 1657 GAGCTATGCTCACACAGGCTGCTTATGTCAACCTCTGACCCCTACACGAGCCCAT 1716
QY 61 TCCTTCACACATTTGGAGCTCTCTGAAGCAAGTCTTATTTGGCAGTTCATTGATGAGAG 120
Db 1717 TCCTTCACACATTTGGAGCTCTCTGAAGCAAGTCTTATTTGGCAGTTCATTGATGAGAG 1776
QY 121 TGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAGAGATATATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGCTTATCAACAGCTTTAGAGAGATATATCGTGGCTTCTTTC 1836
QY 181 TGCTGAGGACACATTTGCAAGCAGAGAGATTTCTTAATGATGTGAGAGTGTGAGAGAG 240
Db 1837 TGCTGAGGACACATTTGCAAGCAGAGAGATTTCTTAATGATGTGAGAGTGTGAGAGAG 1896
QY 241 CCAGTTTCACTACTCATGAGGGGTACATGATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1897 CCAGTTTCACTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1956
QY 301 TAATATTCTCAATTTGGAAAGTAAAGCTGATGATGATGATGATGATGATGATGATGATG 360
Db 1957 TAATATTCTCAATTTGGAAAGTAAAGCTGATGATGATGATGATGATGATGATGATGATG 2016
QY 361 AACTGAGTACAGAGCAGATGATCTCTAAATTCAGATGGAATGCTCAGGGTATG 420
Db 2017 AACTGAGTACAGAGCAGATGATCTCTAAATTCAGATGGAATGCTCAGGGTATG 2076
QY 421 TAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAAATGATCTCCAGATCAGAACT 480
Db 2077 TAGCATGGAAGAAACAAAGCAATTTACATAGATTTTAAATGATCTCCAGATCAGAACT 2136
QY 481 GAAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 2137 GAAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2196
QY 541 GCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAGGTGCTTCA 600
Db 2197 GCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATAGGTGCTTCA 2256
QY 601 AAGAGATCTAGAACAGAGCAAGTCAAGGTCAATCTCTCACTCAGATGTTGGTGTAGT 660
Db 2257 AAGAGATCTAGAACAGAGCAAGTCAAGGTCAATCTCTCACTCAGATGTTGGTGTAGT 2316
QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAAGAAACAACTTAAGGTATTGGG 2376
QY 721 AGATCGATGGGAAACATCTGTAGATGAGACAGAGCGCTGGGTCTTTTACAGA--- 777
Db 2377 AGATCGATGGGAAACATCTGTAGATGAGACAGAGCGCTGGGTCTTTTACAGA--- 2436
QY 778 ----- 777
Db 2437 TCATAGATTACTGCAACAGTTCCTCCCTGAGCTGAAAAGTTTCTTGCTGCTGCTTACAGA 2496
QY 778 ----- 777
Db 2497 AGCTGAAACAACTGCCAATGTCTTACAGGATGCTACCCTGTAAGGAAAGGCTCTCTAGAAGA 2556

CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC cytomegalovirus (CMV) promoter and a small polyA signal sequence.
XX
SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;

Query Match 83.0%; Score 1743; DB 24; Length 4848;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

QY 1 GAGCTATGCTACACACAGCGTCTTATGTCTACACCTCTGACCTTACACGAGGCCATT 60
DB |||||
QY 1680 GAGCTATGCTACACACAGCGTCTTATGTCTACACCTCTGACCTTACACGAGGCCATT 1739
DB |||||
QY 61 TCCTTCACAGCATTTGGAAGCTCTTGAACAGAGTCAATTTGGCAGTTCATTTGATGGAGAG 120
DB |||||
QY 1740 TCCTTCACAGCATTTGGAAGCTCTTGAACAGAGTCAATTTGGCAGTTCATTTGATGGAGAG 1799
DB |||||
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAAGTATTTATCGTGGCTTCTTTC 180
DB |||||
QY 1800 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAAGTATTTATCGTGGCTTCTTTC 1859
DB |||||
QY 181 TGCCTGAGGACATTCGACGACCAAGAGAGATTTCTATGATGTGGAAGTGGTCAAGA 240
DB |||||
QY 1860 TGCCTGAGGACATTCGACGACCAAGAGAGATTTCTATGATGTGGAAGTGGTCAAGA 1919
DB |||||
QY 241 CCAGTTTCTACTCATGAGGGTACATGATGATTTGACAGCCCATTCAGGGCCGGGTGG 300
DB |||||
QY 1920 CCAGTTTCTACTCATGAGGGTACATGATGATTTGACAGCCCATTCAGGGCCGGGTGG 1979
DB |||||
QY 301 TAATATCTCAATTTGGGAGTAACTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 360
DB |||||
QY 1980 TAATATCTCAATTTGGGAGTAACTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 2039
DB |||||
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGGATGCTTCAGGGTAGC 420
DB |||||
QY 2040 AACTGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGATGGGATGCTTCAGGGTAGC 2099
DB |||||
QY 421 TAGCATGGAATAACAAGCAATTTATAGAGTTTTATAGGATCTCCAGAACTCAGAACT 480
DB |||||
QY 2100 TAGCATGGAATAACAAGCAATTTATAGAGTTTTATAGGATCTCCAGAACTCAGAACT 2159
DB |||||
QY 481 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGAAACAGGAAATGGAGGAAGA 540
DB |||||
QY 2160 GAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGAAACAGGAAATGGAGGAAGA 2219
DB |||||
QY 541 GCCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATCAAGTGTCTTCA 600
DB |||||
QY 2220 GCCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTACAAACATCAAGTGTCTTCA 2279
DB |||||
QY 601 AGAGATCTAGAACAGAACTGAGTCAAGTCAAGTCAATTTCTCACTCACTGATGATGATG 660
DB |||||
QY 2280 AGAGATCTAGAACAGAACTGAGTCAAGTCAAGTCAATTTCTCACTCACTGATGATGATG 2339
DB |||||
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACTTAAAGTATTTGGG 720
DB |||||
QY 2340 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACTTAAAGTATTTGGG 2399
DB |||||
QY 721 AGATCGATGGCAACATCTGATAGTGGACAGAGACCGCTGGTCTTTTACAAGA --- 777
DB |||||
QY 2400 AGATCGATGGCAACATCTGATAGTGGACAGAGACCGCTGGTCTTTTACAAGACAC 2459
DB |||||
QY 778 ----- 777
DB 2460 TCATAGATTACTGCAACAGATTCCCTGGACCTGGAAAAAGTTTCTGCTGGGTTACAGA 2519

QY 778 ----- 777
DB 2520 AGCTGAAACAACTGCGCAATGCTCTACAGGATGCTACCCGTAAGGAAAGGCTCTCTAGAAGA 2579
QY 778 ----- 777
DB 2580 CTCAAGGAGTAAAGAGCTGATGAACAATGGAAGACCTCCAAAGGTGAATTTGAAGC 2639
QY 778 ----- 777
DB 2640 TCACACAGATGTTTATCAACACCTGGATGAAACAGCCCAAAAAATCCTGAGATCCCTGGA 2699
QY 778 ----- 777
DB 2700 AGCTTCGATGATGCGAGTCTCTTTACAAAGACGTTTGGATATACATGAACCTTCAAGTGGAG 2759
QY 778 ----- CAGTTCTGACCAAGTG 792
DB 2760 TGAACCTTCGAAAAAGCTCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTTGACCAAGTG 2819
QY 793 GAAGCGTCTGCACCTTTCTCTGACGGAACCTTCTGCTGCTGCTACAGCTGAAAGATGATGA 852
DB 2820 GAAGCGTCTGCACCTTTCTCTGACGGAACCTTCTGCTGCTGCTACAGCTGAAAGATGATGA 2879
QY 853 ATTAAGCGCGCAGGACCTTATTGAGGCGCACTTTCAGCAGATTTCAGAGCAGAACGATCT 912
DB 2880 ATTAAGCGCGCAGGACCTTATTGAGGCGCACTTTCAGCAGATTTCAGAGCAGAACGATCT 2939
QY 913 ACATAGGCGCTTCAAGAGGGAATTTGAAACCTTAAGAACCTGTAATCATGAGTACTCTTGA 972
DB 2940 ACATAGGCGCTTCAAGAGGGAATTTGAAACCTTAAGAACCTGTAATCATGAGTACTCTTGA 2999
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGACCTAGAGAAACCTTACCAGGA 1032
DB 3000 GACTGTACGAATATTTCTGACAGAGCAGCTTTGGAAGGACCTAGAGAAACCTTACCAGGA 3059
QY 1033 GCCCAGAGAGCTGCTCTCTGAGGAGAGCCCAAGAAATGCACTCGGCTTCTACGAAGCA 1092
DB 3060 GCCCAGAGAGCTGCTCTCTGAGGAGAGCCCAAGAAATGCACTCGGCTTCTACGAAGCA 3119
QY 1093 GGCTGAGGAGGTCAATACCTGAGTGGGAAATTTGAACCTGCACTCCGCTGACCTGGCAGAG 1152
DB 3120 GGCTGAGGAGGTCAATACCTGAGTGGGAAATTTGAACCTGCACTCCGCTGACCTGGCAGAG 3179
QY 1153 AAAAATAGATGAGACCCCTTGAAGACTTCCAGGAACTTCAAGAGGCCACGATGAGCTGGA 1212
DB 3180 AAAAATAGATGAGACCCCTTGAAGACTTCCAGGAACTTCAAGAGGCCACGATGAGCTGGA 3239
QY 1213 CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGATCTTGGCAGCCCTGGGCGATCTCCT 1272
DB 3240 CCTCAAGCTGGCCCAAGCTGAGGTGATCAAGGATCTTGGCAGCCCTGGGCGATCTCCT 3299
QY 1273 CATTTGACTCTCTCCAAGATCACTTCCGAGAAAGTCAAGGCACTTTCGAGGAGAAATTTGGCC 1332
DB 3300 CATTTGACTCTCTCCAAGATCACTTCCGAGAAAGTCAAGGCACTTTCGAGGAGAAATTTGGCC 3359
QY 1333 TCTGAAAGAGAACTGAGGCCACGTCATGACCTTGTCTGCGCAGCTTACCCTTTGGGCT 1392
DB 3360 TCTGAAAGAGAACTGAGGCCACGTCATGACCTTGTCTGCGCAGCTTACCCTTTGGGCT 3419
QY 1393 TCAGCTCTCACCGTATTAACCTTCAGCACTCTGGAAGACCTGAAACACACAGATGGAAGCTTCT 1452
DB 3420 TCAGCTCTCACCGTATTAACCTTCAGCACTCTGGAAGACCTGAAACACACAGATGGAAGCTTCT 3479
QY 1453 GCAGTGGCCGCTCGAGGACCGGATGAGGCGAGTGCATGAAAGCCCAAGGACCTTTGGTCC 1512
DB 3480 GCAGTGGCCGCTCGAGGACCGGATGAGGCGAGTGCATGAAAGCCCAAGGACCTTTGGTCC 3539
QY 1513 AGCATCTCAGACACTTTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC 1572
DB 3540 AGCATCTCAGACACTTTCTTCCAGCTCTGTCAGGGTCCCTGGGAGAGAGCCATCTCGCC 3599
QY 1573 AAACAAAGTGGCCCTACTATATCAACACAGAGACTCAAAACAACTTGTGCGGACCATCTCCAA 1632

Db 3600 AACAAGTGCCTACTATATACACAGAGACTCAACAACTTGTGGACATCCCA 3659
Qy 1633 AATGACAGAGCTTACACAGCTTTTACGCTGACCTGAATATGTCAGATTCAGCTATAG 1692
Db 3660 AATGACAGAGCTTACACAGCTTTTACGCTGACCTGAATATGTCAGATTCAGCTATAG 3719
Qy 1693 GACTGCCATGAACCTCGAAGACTGCAAGAGGCTTTGCTGATCTCTTGAGCTGTC 1752
Db 3720 GACTGCCATGAACCTCGAAGACTGCAAGAGGCTTTGCTGATCTCTTGAGCTGTC 3779
Qy 1753 AGCTGCATGTGATCCCTTGACAGCAACACCTCAAGCAAAATGACAGCCCATGATAT 1812
Db 3780 AGCTGCATGTGATCCCTTGACAGCAACACCTCAAGCAAAATGACAGCCCATGATAT 3839
Qy 1813 CCTGCAGATTAATTAATGTTTGACCACTATTATGACCGCTCGAGCAAGACACACAA 1872
Db 3840 CCTGCAGATTAATTAATGTTTGACCACTATTATGACCGCTCGAGCAAGACACACAA 3899
Qy 1873 TTTGGTCAACCTCCTCTCTGCTGGATATGTCGAACTGGCTGCTGAATTTATGA 1932
Db 3900 TTTGGTCAACCTCCTCTCTGCTGGATATGTCGAACTGGCTGCTGAATTTATGA 3959
Qy 1933 TACGGAGCAACAGGAGATCCCTGCTGCTCTTTTAAACTGGCATATTTCCCTGTG 1992
Db 3960 TACGGAGCAACAGGAGATCCCTGCTGCTCTTTTAAACTGGCATATTTCCCTGTG 4019
Qy 1993 TAAAGCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 4020 TAAAGCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4079
Qy 2053 ATTTTGTGACAGCGAGCTGGGCTCTCTGCTGATTTCTATCCAA 2101
Db 4080 ATTTTGTGACAGCGAGCTGGGCTCTCTGCTGATTTCTATCCAA 4128

RESULT 9
AAD37264
ID AAD37264 standard; DNA; 5060 BP.
AC AAD37264;
XX
XX
XX 21-AUG-2002 (first entry)
XX
XX Adeno-associated virus (AAV) vector plasmid, AAV-B-CMV-3849.
XX
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
XX Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
XX WO200183695-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13677.
XX
XX 28-APR-2000; 2000US-200777P.
XX
XX (XIRO/) XIAO X.
XX
XX Xiao X;
XX
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX

PS Example 1; Page 70-71; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter
CC and a small polyA signal sequence.
XX
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;

Query Match 83.0%; Score 1743; DB 24; Length 5060;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

Qy 1 GAGCTATGCTTACACACAGAGCTGCTTATGTCACCACTCTGACCTTACCGAGCCCAT 60
Db 1892 GAGCTATGCTTACACACAGAGCTGCTTATGTCACCACTCTGACCTTACCGAGCCCAT 1951
Qy 61 TCTTTCACAGATTTGGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTCATTTGATGGAGAG 120
Db 1952 TCTTTCACAGATTTGGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTCATTTGATGGAGAG 2011
Qy 121 TGAAGTAAACCTGGACCTGCTTATCAACAGCTTTAGAGAAATTTATCTGGCTTCTTTC 180
Db 2012 TGAAGTAAACCTGGACCTGCTTATCAACAGCTTTAGAGAAATTTATCTGGCTTCTTTC 2071
Qy 181 TCTCAGGACACATTTGGAAGCAAGGAGAGATTTCTAATGATGGAAGTGGTGAAGA 240
Db 2072 TCTCAGGACACATTTGGAAGCAAGGAGAGATTTCTAATGATGGAAGTGGTGAAGA 2131
Qy 241 CCAGTTTCTACTCTCATGAGGGGTACATGATGATTTGACAGCCATCAGGCGCGGTTGG 300
Db 2132 CCAGTTTCTACTCTCATGAGGGGTACATGATGATTTGACAGCCATCAGGCGCGGTTGG 2191
Qy 301 TAATATTTTACAATTTGGAAGTAACTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 2192 TAATATTTTACAATTTGGAAGTAACTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 2251
Qy 361 AACTGAGTACAGAGCAGATCAATCTCTAAATTTCAAGATGGGAATGCCCTCAGGTTAGC 420
Db 2252 AACTGAGTACAGAGCAGATCAATCTCTAAATTTCAAGATGGGAATGCCCTCAGGTTAGC 2311
Qy 421 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAACT 480
Db 2312 TAGCATGGAATAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAACTCAGAACT 2371
Qy 481 GAAAGATTGAATGATGCTGCTTAACAAACAGAGAAAGAACAGGAAATTTGGAGGAAGA 540
Db 2372 GAAAGATTGAATGATGCTGCTTAACAAACAGAGAAAGAACAGGAAATTTGGAGGAAGA 2431
Qy 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTACAAACATAAGGTCCTTCA 600
Db 2432 GCCTCTTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTACAAACATAAGGTCCTTCA 2491
Qy 601 AGAAGATCTAGAAACAAAGTCAAGTCAAGGTCATTTCTCTCACTCAGATGGTGGTGGTAGT 660
Db 2492 AGAAGATCTAGAAACAAAGTCAAGTCAAGGTCATTTCTCTCACTCAGATGGTGGTGGTAGT 2551
Qy 661 TGATGAATCTAGTGGAGTCAAGCAACTGCTTTTGAAGAAACAACTTAAGGTTATGGG 720
Db 2552 TGATGAATCTAGTGGAGTCAAGCAACTGCTTTTGAAGAAACAACTTAAGGTTATGGG 2611
Qy 721 AGATCGATGGCAAAACATCTGTAGATGGAAGAACCGCTGGGTTCTTTTACAAGA --- 777
Db 2612 AGATCGATGGCAAAACATCTGTAGATGGAAGAACCGCTGGGTTCTTTTACAAGACAC 2671

dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene.

Example 1: Page 65-66; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a muscle creatine kinase (MCK) promoter and a small polyA signal sequence.

Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 other;

Query Match 78.1%; Score 1641; DB 24; Length 4414;

Best Local Similarity 88.0%; Pred. No. 0;

Matches 1848; Conservative 0; Mismatches 190; Indels 63; Gaps 3;

Qy	1	GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACCGGAGCCCAT	60
Db	1657	GAGCTATGCCCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACCGGAGCCCAT	1716
Qy	61	TCCTTTCACAGCATTTGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTTCATTTGATGAGAG	120
Db	1717	TCCTTTCACAGCATTTGAAGCTCTGGAAGCAAGTCAATTTGGCAGTTTCATTTGATGAGAG	1776
Qy	121	TGAAGTAAACCTGGACCGTATATCAACACAGCTTTAGAAAGAGTATATCGTGGCTTCTTTC	180
Db	1777	TGAAGTAAACCTGGACCGTATATCAACACAGCTTTAGAAAGAGTATATCGTGGCTTCTTTC	1836
Qy	181	TGCTGAGGACACATTTGAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA	240
Db	1837	TGCTGAGGACACATTTGAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA	1896
Qy	241	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTGG	300
Db	1897	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGTGG	1956
Qy	301	TAATATTTCAATTTGGAAGTAAAGTATGATGATGATGATGATGATGATGATGATGATG	360
Db	1957	TAATATTTCAATTTGGAAGTAAAGTATGATGATGATGATGATGATGATGATGATGATG	2016
Qy	361	AACCTGAAGTACAAGACAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC	420
Db	2017	AACCTGAAGTACAAGACAGATGAATCTCCTAAATTCAGATGGGAATGCCCTCAGGGTAGC	2076
Qy	421	TAGCATGGAAGAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT	480
Db	2077	TAGCATGGAAGAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGAAATCAGAAACT	2136
Qy	481	GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAAGAAAGAAAGAAAGAAAGAA	540
Db	2137	GAAAGAGTTGAATGACTGGCTTAAACAAAACAGAAAGAAAGAAAGAAAGAAAGAAAGAA	2196
Qy	541	GCCTCTTGACCTGATCTTGAAGACCTTAAACCGCCAGTACAAACATTAAGGTGCTTCA	600
Db	2197	GCCTCTTGACCTGATCTTGAAGACCTTAAACCGCCAGTACAAACATTAAGGTGCTTCA	2256
Qy	601	AGAGATCTAGACAGAGCAAGTCAAGTCAATTTCTCCTACCTCACATGTTGGTGTAGT	660
Db	2257	AGAGATCTAGACAGAGCAAGTCAAGTCAATTTCTCCTACCTCACATGTTGGTGTAGT	2316
Qy	661	TGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGGAGAAACAACTTAAGGTATTGGG	720
Db	2317	TGATGAATCTAGTGAGATCAGCAACTGCTGCTTTGGAGAAACAACTTAAGGTATTGGG	2376

Qy	721	AGATCGATGGGCAACACATCTCTAGATGACAGAAAGCCGCTGGGTTCTTTTACAGACAG	780
Db	2377	AGATCGATGGGCAACACATCTCTAGATGACAGAAAGCCGCTGGGTTCTTTTACAGACAT	2436
Qy	781	TTCTGACACAGTGGAGAGCGTCTGCACCTTTCTCTCAGGAACCTTCTGTGTGGCTACAGCT	840
Db	2437	CTTCTCTCAATGGCAACGCTCTTCTGAAAGACAGTGCCTTTTGTAGTATGCTTTTACAG	2496
Qy	841	GAAAGATGATGAATTAAGCCGCGCAGGACCTTATTTGGAGGCGACTTCCAGCAGTTTCAGAA	900
Db	2497	AAAAGAGATGATGATGAACAGATTCACAACTGGCTTTTAAAGATCAAAATGAATGTT	2556
Qy	901	CGAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAACTTAAGAACCTGTAATCAT	960
Db	2557	ATCAAGTCTTCAAAACCTGGCGCTTTTAAAGAGCGATCTAGAAAAGAAAAGCAATCCAT	2616
Qy	961	GAGTACTCTTCAGACTGTACGAATATTTCTCAGACAGCAGCCTTTTGAAGGACTAGAGAA	1020
Db	2617	GGGCAACTGTGTA-----TTTCACTCAAACAGATCTTCTTCAACTGAAG	2662
Qy	1021	ACTTACACAGAGCCCGACAGAGCTGCCTCTCTGAGAGAGAGCCCGAGATGTCACTCGGCT	1080
Db	2663	ATAAGTCAAGTCAAGCCCGACAGAGCAGCATGGCTGGATTAATTTGCCCGTGTGG--	2720
Qy	1081	TCTACGAAAGCAGCGCTGAGGAGTCAATCTAGTGGGAAAAATTTGAACCTGCCTCCG	1140
Db	2721	-----ATTAATTTAGTCCAAAACCTTGAA-----	2743
Qy	1141	TGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTTCAGAACTTCAAGAGGCCAC	1200
Db	2744	-----AAGAGTACAGCACAGACCCCTTTGAAAGACTTCCAGGAACTTCAAGAGGCCAC	2793
Qy	1201	GGATGAGCTGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGT	1260
Db	2794	GGATGAGCTGACCTCAAGCTCGGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCGT	2853
Qy	1261	GGGCGATCTCTCTCATTGACTCTCTCAAGATCACTCTGAGAAAGTCAAGGCACCTTCGAGG	1320
Db	2854	GGGCGATCTCTCTCATTGACTCTCTCAAGATCACTCTGAGAAAGTCAAGGCACCTTCGAGG	2913
Qy	1321	AGAAATTTGGCTCTGAAAGAGAACGTGAGCCAGTCAATGACCTTGTCTGCCAGCTTAC	1380
Db	2914	AGAAATTTGGCTCTGAAAGAGAACGTGAGCCAGTCAATGACCTTGTCTGCCAGCTTAC	2973
Qy	1381	CACCTTTGGGCTTCAAGCTCTCACCGTATAACCTCAGACACTCTGGAAGACCTTGAACACAG	1440
Db	2974	CACCTTTGGGCTTCAAGCTCTCACCGTATAACCTCAGACACTCTGGAAGACCTTGAACACAG	3033
Qy	1441	ATGGAAGCTTTGCAAGTGGCGCTCGAGGACCGAGTCAAGGAGTGCATGAAGCCCAAG	1500
Db	3034	ATGGAAGCTTTGCAAGTGGCGCTCGAGGACCGAGTCAAGGAGTGCATGAAGCCCAAG	3093
Qy	1501	GGACTTTGGTCCAGCATCTCAGCACTTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAG	1560
Db	3094	GGACTTTGGTCCAGCATCTCAGCACTTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAG	3153
Qy	1561	AGCCATCTCGCCCAACAAAGTGCCCTACTATATCAACACAGACTCAACAACTTGTCTG	1620
Db	3154	AGCCATCTCGCCCAACAAAGTGCCCTACTATATCAACACAGACTCAACAACTTGTCTG	3213
Qy	1621	GGACCATCCCAAAATGACAGAGCTTACAGCTTTAGCTGACCTGATATATGTCAGATT	1680
Db	3214	GGACCATCCCAAAATGACAGAGCTTACAGCTTTAGCTGACCTGATATATGTCAGATT	3273
Qy	1681	CTCAGCTTATAGGACTGCGCATGAAACTCCGAGAGACTGCGAAGGCCCTTTGCTTGGATCT	1740
Db	3274	CTCAGCTTATAGGACTGCGCATGAAACTCCGAGAGACTGCGAAGGCCCTTTGCTTGGATCT	3333
Qy	1741	CTTGAGCTGTGAGTGTGATGCTTGGACGACCAACCTCAAGCAAAATGACCA	1800
Db	3334	CTTGAGCTGTGAGTGTGATGCTTGGACGACCAACCTCAAGCAAAATGACCA	3393
Qy	1801	GCCCATGGATATCTCTGCAGATTATTAAATTGTTTGACCACCTATTTATGACCGCTGGAGCA	1860

XX Example 1; Page 53-54; 71pp; English.

CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene Gelta3447 containing nucleotides 1-1992 (N-terminus,
CC hinge H1 and rods R1, R2 and R3), 8749-10227 (rod R24, hinge H4 and

QY 1408 TAACTCAGCACTCTGGAGACCTGAAACACAGATGGAGCTTCTGCAGGTGCGCTGCA 1467
Db 2599 TAACTCAGCACTCTGGAGACCTGAAACACAGATGGAGCTTCTGCAGGTGCGCTGCA 2658
QY 1468 GGACGAGTCAGGACCTGATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACTT 1527
Db 2659 GGACGAGTCAGGACCTGATGAAGCCACACAGGACCTTTGGTCCAGCATCTCAGCACTT 2718
QY 1528 TCTTTCCAGCTCTGCTCAGGCTTCTGGAGAGACCTATCTGCGCAACAAAGTGCCTTA 1587
Db 2719 TCTTTCCAGCTCTGCTCAGGCTTCTGGAGAGACCTATCTGCGCAACAAAGTGCCTTA 2778
QY 1588 CTATATCAACACAGACCTCAACAACTTGTCTGGACCATCCCAAAATGACAGAGCTCTA 1647
Db 2779 CTATATCAACACAGACCTCAACAACTTGTCTGGACCATCCCAAAATGACAGAGCTCTA 2838
QY 1648 CCAGCTCTTTAGTGTGACCTGAATATGTCTGATCTCAGCTTATAGGATGCGCATGAACCT 1707
Db 2839 CCAGCTCTTTAGTGTGACCTGAATATGTCTGATCTCAGCTTATAGGATGCGCATGAACCT 2898
QY 1708 CCGAGACTGCGAGAGGCTTTGCTTGGATCTCTTGAGCTGTGATGCTGATGATGC 1767
Db 2899 CCGAGACTGCGAGAGGCTTTGCTTGGATCTCTTGAGCTGTGATGCTGATGATGC 2958
QY 1768 CTTGGACACGACCAACCTCAAGCAAAATGACACAGCCCATGGATATCTCTGCAATATTAA 1827
Db 2959 CTTGGACACGACCAACCTCAAGCAAAATGACACAGCCCATGGATATCTCTGCAATATTAA 3018
QY 1828 TTGTTTGACCACTATTATGACCGCTGCGAGCAAGACACACAAATTTGGTCAAGCTCC 1887
Db 3019 TTGTTTGACCACTATTATGACCGCTGCGAGCAAGACACACAAATTTGGTCAAGCTCC 3078
QY 1888 TCTCTCGCTGGATGTGCTGAACTGAGCTGTGCTGATTTTATGATAGCGGACGAAACAGG 1947
Db 3079 TCTCTCGCTGGATGTGCTGAACTGAGCTGTGCTGATTTTATGATAGCGGACGAAACAGG 3138
QY 1948 GAGATCCGCTGCTCTTTTAACTGGCATCTTCCCTGTGTAAGCAATTTGTGACCAAGC 2007
Db 3139 GAGATCCGCTGCTCTTTTAACTGGCATCTTCCCTGTGTAAGCAATTTGTGACCAAGC 3198
QY 2008 AGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGC 2067
Db 3199 AGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGC 3258
QY 2068 CAGGCTGGGCTCTCTCTGATGATTTATCTATCCAA 2101
Db 3259 CAGGCTGGGCTCTCTCTGATGATTTATCTATCCAA 3292

RESULT 14

AAD37234
ID AAD37234 standard; DNA; 3999 BP.

XX AC

XX AC

XX AC

DT 21-AUG-2002 (first entry)

XX Human dystrophin minigene delta3990.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;

KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;

KW Becker muscular dystrophy; ds.

XX OS

XX Homo sapiens.

XX WO2001:83695-A2.

XX PD

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-20077P.

XX 28-APR-2000; 2000US-20077P.

XX (XIAO/) XIAO X.

XX Xiao X;

XX WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophy comprises an N-terminal domain or modified N-terminal domain,

PT rod repeats, H1 and H4 domains and a cysteine rich domain of a

PT dystrophin gene -

XX Example 1; Page 46-47; 71pp; English.

XX The present invention relates to an isolated nucleotide sequence encoding

CC a dystrophin minigene. The minigene comprises N-terminal or modified

CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4

CC domains and cysteine-rich domains of dystrophin or utrophin genes. The

CC invention also relates to a recombinant adeno-associated virus (AAV)

CC comprising dystrophin minigene operably linked to an expression control

CC element, in a recombinant adeno-associated virus or retrovirus is

CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular

CC dystrophy (BMD) in a mammalian subject. The present sequence is human

CC dystrophin minigene delta3990 containing nucleotides 1-1668 (N-terminus,

CC hinge H1 and rods R1 and R2), 8059-10227 (rods R22, R23 and R24, hinge

CC H4 and CR domain) and 11047-11058 (dystrophin last 3 amino acids).

XX Sequence 3999 BP; 1223 A; 907 C; 933 G; 936 T; 0 other;

Query Match

Best Local Similarity 76.2%; Score 1602; DB 24; Length 3999;

Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;

QY 1 GAGCTATGCTTACACACAGGCTCTTATGTCACACCTCTGACCTTACAGGAGCCCAT 60

Db 900 GAGCTATGCTTACACACAGGCTCTTATGTCACACCTCTGACCTTACAGGAGCCCAT 959

QY 61 TCCTTACACATTTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTTCAATTTGATGGAGAG 120

Db 960 TCCTTACACATTTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTTCAATTTGATGGAGAG 1019

QY 121 TGAAGTAAACCTGGACCGCTTATCAAAAGTATTTAGAGAGTATTTATCGTGGCTTCTTTC 180

Db 1020 TGAAGTAAACCTGGACCGCTTATCAAAAGTATTTAGAGAGTATTTATCGTGGCTTCTTTC 1079

QY 181 TGCTGAGGACACATTGCAACACAGGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA 240

Db 1080 TGCTGAGGACACATTGCAACACAGGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA 1139

QY 241 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCGCGGTGG 300

Db 1140 CCAGTTTTCATCTCATGAGGGGTACATGATGGATTGACAGCCCATCAGGCGCGGTGG 1199

QY 301 TAATATTCTACAAATGGGAAGTAACTGATGGAAACAGGAAATTTATCAGAAGATGAAGA 360

Db 1200 TAATATTCTACAAATGGGAAGTAACTGATGGAAACAGGAAATTTATCAGAAGATGAAGA 1259

QY 361 AACTGAAAGTACAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCTCTCAGGGTAGC 420

Db 1260 AACTGAAAGTACAGAGCAGATGAATCTCTAAATTTCAAGATGGGAATGCTCTCAGGGTAGC 1319

QY 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAACT 480

Db 1320 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGATCAGAACT 1379

QY 481 GAAAGAGTTGAATGACTGGCTTAAACAAAAACAGAGAGAAAGAACAGGAAAAATGGAGGAAGA 540

Db 1380 GAAAGAGTTGAATGACTGGCTTAAACAAAAACAGAGAGAAAGAACAGGAAAAATGGAGGAAGA 1439

QY 541 GCCTCTCGACCTGATCTTGAAGACCTTAAACGCGCAAGTACAAACATAGGCTCTTCA 600

Db 1440 GCCTCTCGACCTGATCTTGAAGACCTTAAACGCGCAAGTACAAACATAGGCTCTTCA 1499

501 AGAGATCTAGACAGAAAGCTCAGGCTCAATTTCTCTCACTCACTATGGTGGTGTAGT 660
1500 AGAAGATCTAGAACAGAAAGCTCAGGCTCAATTTCTCTCACTCACTATGGTGGTGTAGT 1559
661 TGAATGAATCTAGTGGAGATCAGCAACTCTCTGTTTGGAGAACAACTTAAAGTATTGGG 720
1560 TGAATGAATCTAGTGGAGATCAGCAACTCTCTGTTTGGAGAACAACTTAAAGTATTGGG 1619
721 AGATCATGGGCAAAATCTGTAGATGGACAGAACCGCTGGTCTTTTACAGA --- 777
1620 AGATCATGGGCAAAATCTGTAGATGGACAGAACCGCTGGTCTTTTACAGAACA 1679
778 ----- 777
1680 GCCTGACCTAGCTCTCTGGACTGACCACTATTGGAGCTCTCTCTACTCAGACTGTTACTCT 1739
778 ----- 777
1740 GGTGACACAACCTGTGGTTACTAAGGAACTGCCATCTCTCCAACTAGAAATGCCATCTTC 1799
778 ----- 777
1800 CTTGATGTTGGAGTACTACTCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAA 1859
778 ----- 777
1860 GTTCTTCTGCTGCTTACAGAGCTGAAACAACTGCCAATGTCTCTACAGGATGCTACCG 1919
778 ----- 777
1920 TAAGGAAAGCTCTCTAGAGACTCCAAAGGATGAAAGAGCTGATGAAACAATGGCAAGA 1979
778 ----- 777
1980 CCTCAGGTGAAATTGAAGCTCACAGAGTGTATTATCAACCTGGATGAAAGCAAGCA 2039
778 ----- 777
2040 AAAAATCTCAGATCCCTGGAAGGTTCCGATGATGAGTCTGTTACAAAGAGCTTTGA 2099
778 ----- 777
2100 TAACATGAATCTCAAGTGGAGTGAATCTCGGAAAAGTCTCTCAACATTAGTCCCATTT 2159
778 -----CAGTTCTGAACAGTGGAGCGTGTGACCTTTCTCTGACGAACTTCTGTTG 831
2160 GGAAGCAGTCTCTGACAGTGGAGCGTCTGCACTTTCTCTGACGAACTTCTGTTG 2219
832 GCTACAGCTCAAGAGATGATGAATTAAGCGGCGAGGACCTATTGGAGGCGACTTTCCAGC 891
2220 GCTACAGCTCAAGAGATGATGAATTAAGCGGCGAGGACCTATTGGAGGCGACTTTCCAGC 2279
892 AGTTCAAGAGCAGAACGATCTAATAGGCGCTTCAAGAGGGAATTAAGAACTAAGAAC 951
2280 AGTTCAAGAGCAGAACGATCTAATAGGCGCTTCAAGAGGGAATTAAGAACTAAGAAC 2339
952 TGTATCATGAGTACTCTTGAAGCTGTACAAATTTCTGACAGAGCGCTTTGGAAG 1011
2340 TGTATCATGAGTACTCTTGAAGCTGTACAAATTTCTGACAGAGCGCTTTGGAAG 2399
1012 ACTAGAGAACTCTACAGAGCGCCAGAGAGCTGCTCTCTGAGAGAGAGCCAGAACT 1071
2400 ACTAGAGAACTCTACAGAGCGCCAGAGAGCTGCTCTCTGAGAGAGAGCCAGAACT 2459
1072 CACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAATTTGAACT 1131
2460 CACTCGGCTTCTAGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAAATTTGAACT 2519
1132 GCATCTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAATTC 1191
2520 GCATCTCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCCAGGAATTC 2579

QY 1192 AGAGGCCACGATGAGCTGAGCCTCAAGCTGCGCCAAAGCTGAGTGAATCAAGGATCCTG 1251
Db 2580 AGAGGCCACGATGAGCTGAGCCTCAAGCTGCGCCAAAGCTGAGTGAATCAAGGATCCTG 2639
QY 1252 GCAGCCGCTGGGATCTCTCTATTGACTCTCTCAAGATCACTCGAGAAAGTCAAGGC 1311
Db 2640 GCAGCCGCTGGGATCTCTCTATTGACTCTCTCAAGATCACTCGAGAAAGTCAAGGC 2699
QY 1312 ACTTCGAGGAGAAATTCGGCTCTGAAAGAGAAAGCTGAGCCAGCTCAATGACCTTGTCTG 1371
Db 2700 ACTTCGAGGAGAAATTCGGCTCTGAAAGAGAAAGCTGAGCCAGCTCAATGACCTTGTCTG 2759
QY 1372 CCAGCTTACCACTTTTGGGATTCAGCTCTCAGCTATACCTCAGCAGCTCTGAGAGCT 1431
Db 2760 CCAGCTTACCACTTTTGGGATTCAGCTCTCAGCTATACCTCAGCAGCTCTGAGAGCT 2819
QY 1432 GAACACAGAGTGAAGCTTCTGAGGCTGGCGCTCGAGGACCGAGTCAAGCAGCTGCATGA 1491
Db 2820 GAACACAGAGTGAAGCTTCTGAGGCTGGCGCTCGAGGACCGAGTCAAGCAGCTGCATGA 2879
QY 1492 AGCCCAAGGAGCTTTGGTCCAGCATCTCAGCATCTTCTTCCAGCTCTCTCCAGGCTCC 1551
Db 2880 AGCCCAAGGAGCTTTGGTCCAGCATCTCAGCATCTTCTTCCAGCTCTCTCCAGGCTCC 2939
QY 1552 CTGGGAGAGAGCTTCGCGCAACAAAGTGCCTACTATCAACCCAGAGACTCAAA 1611
Db 2940 CTGGGAGAGAGCTTCGCGCAACAAAGTGCCTACTATCAACCCAGAGACTCAAA 2999
QY 1612 AACTTGTGTGGAGCCATCCCAAAATGACAGAGCTCTACAGCTTTTAGCTCACTGAATAA 1671
Db 3000 AACTTGTGTGGAGCCATCCCAAAATGACAGAGCTCTACAGCTTTTAGCTCACTGAATAA 3059
QY 1672 TGTGAGATTTCTAGCTTATAGGACTGCCATGAACCTCCGAGACTCGAGAGGCCCTTTG 1731
Db 3060 TGTGAGATTTCTAGCTTATAGGACTGCCATGAACCTCCGAGACTCGAGAGGCCCTTTG 3119
QY 1732 CTGAGATCTCTTGGAGCTGTGAGTGTGATGCTTGGACAGCAGCAACCTCAAGCA 1791
Db 3120 CTGAGATCTCTTGGAGCTGTGAGTGTGATGCTTGGACAGCAGCAACCTCAAGCA 3179
QY 1792 AAATGAACAGCCGATGATATCTGAGATTTAATGTTTGAACCACTATTTATGACCG 1851
Db 3180 AAATGAACAGCCGATGATATCTGAGATTTAATGTTTGAACCACTATTTATGACCG 3239
QY 1852 CCTGAGCAAGAGCAGCAACAAATTTGGTCAACGCTCTCTGGTGGATATGTCTGAA 1911
Db 3240 CCTGAGCAAGAGCAGCAACAAATTTGGTCAACGCTCTCTGGTGGATATGTCTGAA 3299
QY 1912 CTGCTCTGATGATTTATGATACGGAACGAAACAGGAGGATCGGTCTCTGCTTTAA 1971
Db 3300 CTGCTCTGATGATTTATGATACGGAACGAAACAGGAGGATCGGTCTCTGCTTTAA 3359
QY 1972 AACTGGCATCAATTTCCCTGTGTAAGCACAATTTGGAAGACAAGTACAGATACCTTTCAA 2031
Db 3360 AACTGGCATCAATTTCCCTGTGTAAGCACAATTTGGAAGACAAGTACAGATACCTTTCAA 3419
QY 2032 GCAAGTGGCAAGTTCAACAGGATTTGTGACAGCGCAGGCTGGGCTCTCTCTGCAATGA 2091
Db 3420 GCAAGTGGCAAGTTCAACAGGATTTGTGACAGCGCAGGCTGGGCTCTCTCTGCAATGA 3479
QY 2092 TTTATCCAA 2101
Db 3480 TTTATCCAA 3489

RESULT 15

AAD37256

ID AAD37256 standard; DNA; 4966 BP.

XX

AC AAD37256;

XX

DT 21-AUG-2002 (first entry)

XX

Adeno-associated virus vector plasmid, AAV-MCK-delta3990.

Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.

Chimeric - Homo sapiens.

Chimeric - Unidentified.

WO200193695-A2.

08-NOV-2001.

27-APR-2001; 2001WO-US13677.

28-APR-2000; 2000US-200777P.

(XIAO/) XIAO X.

Xiao X;

WPI; 2002-049342/06.

New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a
dystrophin gene.

Example 1; Page 59-60; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified
N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is AAV
vector plasmid construct containing human dystrophin minigenes, a muscle
creatine kinase (MCK) promoter and a small polyA signal sequence.

Sequence 4966 BP; 1403 A; 1216 C; 1232 G; 1115 T; 0 other;

Query Match 76.2%; Score 1602; DB 24; Length 4966;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;

1 GAGCTATGCTACACACAGGCTGCTATGTCACCACTCTGACCCCTACACGAGCCATT 60
1657 GAGCTATGCTACACACAGGCTGCTATGTCACCACTCTGACCCCTACACGAGCCATT 1716

61 TCCTTTCACAGCATTTGGAAGCTCTGGAAGCAAGTCATTGGCAGTTCATTGATGAGAG 120

1717 TCCTTTCACAGCATTTGGAAGCTCTGGAAGCAAGTCATTGGCAGTTCATTGATGAGAG 1776

121 TGAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAGTATTATCGTGGTCTTTC 180

1777 TGAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAGTATTATCGTGGTCTTTC 1836

181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240

1837 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1896

241 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACGCCCATCAGGCCCGGGTTGG 300

1897 CCAGTTTCATCTCATGAGGGGTACATGATGATTTGACGCCCATCAGGCCCGGGTTGG 1956

301 TAATATTTCTCAATTTGGAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 360

1957 TAATATTTCTCAATTTGGAGTAAAGCTGATTGGAAACAGGAAATTTATCAGAAGATGAAGA 2016

361 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCCTCAGGTAGC 420
2017 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCCTCAGGTAGC 2076
421 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAACTCAGAAACT 480
2077 TAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAACTCAGAAACT 2136
481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGAAAAATGGAGGAGA 540
2137 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAGAAAAATGGAGGAGA 2196
541 GCCTCTTGACCTGATCTTTGAAGACCTTAAAGCGCAAGTACAAACAACATTAAGGTGTTCA 600
2197 GCCTCTTGACCTGATCTTTGAAGACCTTAAAGCGCAAGTACAAACAACATTAAGGTGTTCA 2256
601 AGAAGATCTAGAACAAAGTCAAGGTCAAGTCTCTCACTCACTCACTCACTCACTCACTCACT 560
2257 AGAAGATCTAGAACAAAGTCAAGGTCAAGTCTCTCACTCACTCACTCACTCACTCACTCACT 2316
661 TGATGAATCTAGTGGAGATCAGCACTGCTGTTTGGAGAAACAACATTAAGGTATTGGG 720
2317 TGATGAATCTAGTGGAGATCAGCACTGCTGTTTGGAGAAACAACATTAAGGTATTGGG 2376
721 AGATCGATGGGCAACATCTGTAGATGGACAGACCGCTGGTTCTTTTACAAGA --- 777
2377 AGATCGATGGGCAACATCTGTAGATGGACAGACCGCTGGTTCTTTTACAAGA --- 2436
778 --- 777
2437 GCCTGACCTAGTCTCTGAGTGAACCACTATTGGAGCCTCTCTCTCTCAGACTGTTACTCT 2496
778 --- 777
2497 GGTGACACAACTGTGGTTACTTAAGAAACTGCCATCTCCAACTAGAAATGCCATCTTC 2556
778 --- 777
2557 CTTGATGTTGGAGGTACTTACTCATAGATTCTGCAACAGTTCCCTCCCTGGACCTGAAAA 2616
778 --- 777
2617 GTTCTTTCCTGCTTACAGAACTGCAACACTGCAATGTCTCTACAGGATGCTACCCG 2676
778 --- 777
2677 TAAGGAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAAACAATGGCAAGA 2736
778 --- 777
2737 CCTCCAAAGTGAAATTAAGACTCACACAGATGTTTATCAACACTGGATGAAAAACAGCCA 2796
778 --- 777
2797 AAAAATCCTGAGATCCCTGGAGGTTCCGATGATGAGTCTCTGTTACAAAGACGTTTGA 2856
778 --- 777
2857 TAACATGAACCTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGGTCCCAATTT 2916
778 --- CAGTTCTGACCACTGGAGCGCTGCACTTTCTCTGCGAGAACTTTCTGGTGTG 831
2917 GGAAGCCAGTTCTGACCACTGGAGCGCTGCACTTTCTCTGCGAGAACTTTCTGGTGTG 2976
832 GCTACAGCTGAAAGATGAAATTAAGCCGAGGCACTTATTGGAGGCGACTTTTCAGC 891
2977 GCTACAGCTGAAAGATGAAATTAAGCCGAGGCACTTATTGGAGGCGACTTTTCAGC 3036
892 AGTTCAAGACAGAACGATGTACATAGGCTTCAAGAGGGAATTTGAAACTTAAGAAACC 951
3037 AGTTCAAGACAGAACGATGTACATAGGCTTCAAGAGGGAATTTGAAACTTAAGAAACC 3096
952 TGTATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGACGCGCTTTTGAAGG 1011

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:11:40 ; Search time 129.879 Seconds
(without alignments)
7140.092 Million cell updates/sec

Title: US-09-845-416-12_COPY_900_3000

Perfect score: 2101
Sequence: 1 GGGCTAGCTACACAGG.....tttgcgattctatccaa 2101

Scoring table: IDENTITY NUC

Gapop 10.0 , Capext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B-COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A-COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B-COMB.seq: *
5: /cgn2_6/prodata/2/ina/PTUS-COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1328.8	63.2	5952	4	US-09-687-875A-1
2	1317.8	62.7	13977	4	US-09-484-970B-60
3	1149.6	54.7	19307	3	US-08-836-022A-10
4	1149.6	54.7	19307	3	US-09-427-048A-10
5	547.4	26.1	6045	4	US-09-091-501B-7
6	547.4	26.1	10320	4	US-09-091-501B-9
7	79.4	3.8	200	4	US-09-091-501B-5
8	78.6	3.7	200	4	US-09-091-501B-4
9	78.6	3.7	200	4	US-09-091-501B-6
10	76.6	3.6	7218	1	US-08-232-463-14
11	44.2	2.1	2574	4	US-09-668-313A-10
12	44	2.1	1230025	4	US-09-198-452A-1
13	43.4	2.1	1179	4	US-09-107-532A-1186
14	42.8	2.0	1690	4	US-09-620-312D-69
15	42.8	2.0	7812	3	US-09-368-590-1
16	40.4	1.9	2223	1	US-08-257-073-4
17	39.2	1.9	16995	4	US-08-961-527-82
18	38.6	1.8	1995	1	US-08-425-069-3
19	38.6	1.8	1995	2	US-08-317-844B-3
20	38.4	1.8	7672	4	US-09-220-132-24
21	38.2	1.8	428	4	US-09-668-313A-3
22	38.2	1.8	1131	6	5180810-3
23	38.2	1.8	1784	6	5180810-2
24	38.2	1.8	4439	4	US-09-668-313A-17
25	38	1.8	1394	4	US-09-247-155-76
26	36.8	1.8	1886	6	5210183-1
27	36.6	1.7	1845	4	US-08-887-534A-22

Sequence 22, Appl
Sequence 17, Appl
Sequence 17, Appl
Sequence 14, Appl
Sequence 1, Appl
Sequence 193, Appl
Sequence 13, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 30, Appl
Sequence 5, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 38, Appl
Sequence 10, Appl
Sequence 4055, Ap
Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-687-875A-1
; Sequence 1, Application US/09687875A
; Patent No. 6544786
; GENERAL INFORMATION:
; APPLICANT: Xiao, Xiao
; APPLICANT: Liu, Paul
; TITLE OF INVENTION: METHOD AND VECTOR FOR PRODUCING AND TRANSFERRING TRANS-SPICED PEI
; FILE REFERENCE: 00792
; CURRENT APPLICATION NUMBER: US/09/687,875A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/458,968
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5952
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2897)..(2898)
; OTHER INFORMATION: S4 junction site
; NAME/KEY: misc feature
; LOCATION: (3198)..(3199)
; OTHER INFORMATION: S2 junction site
US-09-687-875A-1

Query Match 63.2%; Score 1328.8; DB 4; Length 5952;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 762 GGGTCTTTTACAGACAGTTCTGACAGTGAAGCGTCTGCACCTTTCTCTGCAGAAC 821
DB 3284 GGTCCATTGGAGCCAGTTCTGACAGTGAAGCGTCTGCACCTTTCTCTGCAGAAC 3343
QY 822 TTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGGACCTATTGGAGCG 881
DB 3344 TTCTGGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGGACCTATTGGAGCG 3403
QY 882 ACTTTCAGAGTTTCAAGACGAAGATGATCATAGGCGCTTCAAGAGGGAATTCAGAA 941
DB 3404 ACTTTCAGAGTTTCAAGACGAAGATGATCATAGGCGCTTCAAGAGGGAATTCAGAA 3463
QY 942 CTAAAGAACTGTAATCATGAGTACTTCTTGAGACTGTAGAAATTTCTGACAGAGCAG 1001
DB 3464 CTAAAGAACTGTAATCATGAGTACTTCTTGAGACTGTAGAAATTTCTGACAGAGCAG 3523
QY 1002 CTTTGAAGACTAGAGAACTCTACAGAGCCAGAGCTGCTCTCTGAGGAGAGAG 1061

3524 CTTTGAAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAG 3583
1062 CCAGAAATGTCACTGGCTTCTAAGAAAGAGGCTGAGGAGGTCAATCTAGTGGGAAA 1121
3584 CCAGAAATGTCACTGGCTTCTAAGAAAGAGGCTGAGGAGGTCAATCTAGTGGGAAA 3643
1122 AATTGAACCTGCACTCCCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC 1181
3644 AATTGAACCTGCACTCCCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC 3703
1182 AGGAATCTCAAGAGCCACGATGAGTGGACCTCAAGCTCGGCAAGCTGAGGTGATCA 1241
3704 AGGAATCTCAAGAGCCACGATGAGTGGACCTCAAGCTCGGCAAGCTGAGGTGATCA 3763
1242 AGGATCTCGGAGCCGCTGGCGCATCTCCCTCATTTGACTCTCTCAAGATCACTTCGAGA 1301
3764 AGGATCTCGGAGCCGCTGGCGCATCTCCCTCATTTGACTCTCTCAAGATCACTTCGAGA 3823
1302 AAGTCAAGGCACTTCGAGGAGAAAATTGGCCTCTGAAAGAAAGCTGAGCAAGCTCAATG 1361
3824 AAGTCAAGGCACTTCGAGGAGAAAATTGGCCTCTGAAAGAAAGCTGAGCAAGCTCAATG 3883
1362 ACCTTGCTCGCAGTTCACCTTTGGCATTTGAGCTCTCAGCTCTCAGCTATACTCAGACTC 1421
3884 ACCTTGCTCGCAGTTCACCTTTGGCATTTGAGCTCTCAGCTCTCAGCTATACTCAGACTC 3943
1422 TGAAGACCTGAACACCAAGAGAGCACTCTCCGAGGTGGCGCTCGAGGAGCCAGCTCAGGC 1481
3944 TGAAGACCTGAACACCAAGAGAGCACTCTCGAGGTGGCGCTCGAGGAGCCAGCTCAGGC 4003
1482 AGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCATCTTTTTCAGCTCTG 1541
4004 AGCTGCATGAAGCCACAGGACCTTTGGTCCAGCATCTCAGCATCTTTTTCAGCTCTG 4063
1542 TCCAGGGTCCCTGGGAGAGAGCCATCTCCGCAAAACAAAGTCCCTACTATATCAACCAAG 1601
4064 TCCAGGGTCCCTGGGAGAGAGCCATCTCCGCAAAACAAAGTCCCTACTATATCAACCAAG 4123
1602 AGACTCAAAACAACTTCTGGGAGACATCCCAAAATGACAGACTCTACAGCTTTTAGCTG 1661
4124 AGACTCAAAACAACTTCTGGGAGACATCCCAAAATGACAGACTCTACAGCTTTTAGCTG 4183
1662 ACCTGAATATCTCAGATTCTCAGCTTATAGACTGCGCATGAAACTCCGAAAGACTCGAGA 1721
4184 ACCTGAATATCTCAGATTCTCAGCTTATAGACTGCGCATGAAACTCCGAAAGACTCGAGA 4243
1722 AGGCCCTTTGCTTGGATCTCTTGAAGCTGTGAGCTGCAATGATGCTTGGACAGACCA 1781
4244 AGGCCCTTTGCTTGGATCTCTTGAAGCTGTGAGCTGCAATGATGCTTGGACAGACCA 4303
1782 ACCTCAGCAAAATGACAGCCCATGGATATCCCTGAGATTATTAAATGTTTGAACCACTA 1841
4304 ACCTCAGCAAAATGACAGCCCATGGATATCCCTGAGATTATTAAATGTTTGAACCACTA 4363
1842 TTTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGATA 1901
4364 TTTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACGCTCCCTCTCTGCGTGGATA 4423
1902 TGTGTCTGACTGCTGCTGATGTTTATGATACGGGACGACAGGAGGATCCGCTGCC 1961
4424 TGTGTCTGACTGCTGCTGATGTTTATGATACGGGACGACAGGAGGATCCGCTGCC 4483
1962 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTGTAAGACACATTTGGAAGACAAGTACAGAT 2021
4484 TGTCTTTTAAAACTGGCATCATTTCCCTGTGTGTAAGACACATTTGGAAGACAAGTACAGAT 4543
2022 ACCTTTTCAAGCAAGTGGAGTTCACAGGATTTTGTGACAGGCGAGGCTGGGCTCC 2081
4544 ACCTTTTCAAGCAAGTGGAGTTCACAGGATTTTGTGACAGGCGAGGCTGGGCTCC 4603
2082 TTTCTGATCATTTCTATCCAA 2101
4604 TTTCTGATCATTTCTATCCAA 4623

RESULT 2

US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CB1
; NAME/KEY: unsure
; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
; US-09-484-970B-60

Query Match 62.7%; Score 1317.8; DB 4; Length 13977;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1333; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 762 GGGTCTTTTAAAGACAGTCTGACAGTGGAGCGTCTGACCTTTCTTCTGAGGAGAC 821
DB 8598 GGTCCCAATTTGAAGCCAGTCTGACAGTGGAGCGTCTGACCTTTCTTCTGAGGAGAC 8657
QY 822 TTCTGTGTGGCTAGAGTGAAGATGATGAATTAAGCGCGGAGGACCTATTGAGGCG 881
DB 8658 TTCTGTGTGGCTAGAGTGAAGATGATGAATTAAGCGCGGAGGACCTATTGAGGCG 8717
QY 882 ATTTTCCAGCACTTCAAGAGCAAGATGATGATGAGGCGCTTCAAGAGGGAATTTGAAA 941
DB 8718 ACTTTCCAGCACTTCAAGAGCAAGATGATGATGAGGCGCTTCAAGAGGGAATTTGAAA 8777
QY 942 CTAAAGAACCTTCAATCATGATGATCTTCTGAGCTGTAGCAATATTCTGACAGAGCAGC 1001
DB 8778 CTAAAGAACCTTCAATCATGATGATCTTCTGAGCTGTAGCAATATTCTGACAGAGCAGC 8837
QY 1002 CTTTGGAGAGCTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAG 1061
DB 8838 CTTTGGAGAGCTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAG 8897
QY 1062 CCCAGATGTCTACTGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAA 1121
DB 8898 CCCAGATGTCTACTGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAA 8957
QY 1122 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC 1181
DB 8958 AATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAAGACTCC 9017
QY 1182 AGGAATCTCAAGAGCCACGATGAGTGGACCTCAGCTGGCGCAAGCTGAGGTGATCA 1241
DB 9018 AGGAATCTCAAGAGCCACGATGAGTGGACCTCAGCTGGCGCAAGCTGAGGTGATCA 9077
QY 1242 AGGATCTCGGAGCCGCTGGCGCATCTCCTCATTTGACTCTCTCCTCAAGATCACTTCGAGA 1301
DB 9078 AGGATCTCGGAGCCGCTGGCGCATCTCCTCATTTGACTCTCTCCTCAAGATCACTTCGAGA 9137
QY 1302 AAGTCAAGCACTTCGAGGAGAAAATTGGCTCTCTGAAAGAGAACTGAGGCGACGCTCAATG 1361
DB 9138 AAGTCAAGCACTTCGAGGAGAAAATTGGCTCTCTGAAAGAGAACTGAGGCGACGCTCAATG 9197
QY 1362 ACCTTGCTCGCAGCTTACCACTTTGGGGCAATTCAGCTCTCACCGTATAACCTTCAGCACTC 1421

Db 9198 ACCTTGCTGCCAGCTTACCACTTTGGGATTCAGCTCTCACCGTATAACCTCAGCACTC 9257
Qy 1422 TGAAGACCTTGAACACACAGATGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCCAGGC 1481
Db 9258 TGAAGACCTTGAACACACAGATGAAGCTTCTGAGGTGGCCGTGAGGACCGAGTCCAGGC 9317
Qy 1482 AGCTGATGAAGCCCAAGGACTTTGGTCCAGACTCTCAGCACTTCTTCCAGTCTG 1541
Db 9318 AGCTGATGAAGCCCAAGGACTTTGGTCCAGACTCTCAGCACTTCTTCCAGTCTG 9377
Qy 1542 TCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 1601
Db 9378 TCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACACG 9437
Qy 1602 AGACTCAACAACTTGTGGGACATCCCAAATGACAGCTCTACAGTCTTTAGCTG 1661
Db 9438 AGACTCAACAACTTGTGGGACATCCCAAATGACAGCTCTACAGTCTTTAGCTG 9497
Qy 1662 ACCTGAATATGTCAATCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGA 1721
Db 9498 ACCTGAATATGTCAATCTCAGCTTATAGGACTGCCATGAACTCCGAAGACTGCAGA 9557
Qy 1722 AGCCCTTTGCTGGATCTCTGAGCTGTGATGCTGATGCTTGGACCAACA 1781
Db 9558 AGCCCTTTGCTGGATCTCTGAGCTGTGATGCTGATGCTTGGACCAACA 9617
Qy 1782 ACCTCAAGCAAAATGACCAAGCCATGATATCTGAGATTTATTTGACCACTA 1841
Db 9618 ACCTCAAGCAAAATGACCAAGCCATGATATCTGAGATTTATTTGACCACTA 9677
Qy 1842 TTTATGACCCCTGGAGAGACACAAATTTGGTCAACCTCTCTCTGCTGGATA 1901
Db 9678 TTTATGACCCCTGGAGAGACACAAATTTGGTCAACCTCTCTCTGCTGGATA 9737
Qy 1902 TGTGCTGAAGTGGCTGCTGAATTTATGATGGGACGACAGGAGGATCCGTGCTC 1961
Db 9738 TGTGCTGAAGTGGCTGCTGAATTTATGATGGGACGACAGGAGGATCCGTGCTC 9797
Qy 1962 TGTCTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAGTACAGAT 2021
Db 9798 TGTCTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAGTACAGAT 9857
Qy 2022 ACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGCTT -GGGCTTC 2080
Db 9858 ACCTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCGAGGCTGGGCTTC 9917
Qy 2081 CTCTCGATGATTTATCCAA 2101
Db 9918 CTCTCGATGATTTATCCAA 9938

RESULT 3

JS-08-836-022A-10/c
Sequence 10, Application US/08836022A
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
APPLICANT: Wilson, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Chen, Shu-Jen
APPLICANT: Weltzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,022A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPVN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-836-022A-10

Query Match 54.7%; Score 1149.6; DB 3; Length 19307;

Best Local Similarity 91.1%; Pred. No. 0;
Matches 1221; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy 762 GGGTCTCTTTTACAAGACAGATTCTGACCAGTGGAGCGTCTGCACCTTTCTCTCAGGAAC 821
Db 6096 GGTCCCATTTGGAAGCAAGTTCTGACCAGTGGAGCGTTTTCATCTTTCTCAGGAAC 6037
Qy 822 TTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCG 881
Db 6036 TTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTATTGGAGCG 5977
Qy 882 ACTTTCAGCAGTTCAGAAGCAGACGATGTACATAGGGGCTTCAAGAGGAAATTGAAA 941
Db 5976 ATTTCAGCAGTTCAGAAGCAGATGATATATAGAGGCTTCAAGAGGAAATTGAAA 5917
Qy 942 CTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTGACAAATTTCTGACAGAGCAG 1001
Db 5916 CTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTGACAAATTTCTGACAGAGCAG 5857
Qy 1002 CTTTGGAGGACTAGAGAACTCTACAGAGCCGACAGAGCTGCCTCTCTGAGAGAGAG 1061
Db 5856 CTTTGGAGGACTAGAGAACTCTACAGAGCCGACAGAACTGCCTCTCTGAGAGAGAG 5797
Qy 1062 CCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACCTGAGTGGGAA 1121
Db 5796 CTCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAACGCTGAATGGACA 5737
Qy 1122 AATTGAACCTGCACTCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAGACTCC 1181
Db 5736 AATTGAACCTGCACTCGCTGACTGGCAGAGAAAATAGATGAGACCTTGAAGACTCC 5677
Qy 1182 AGGAACCTTCAAGAGGCGCACCGATGAGCTGCACTCAAGCTGCGCCCAAGCTGAGGTGATCA 1241
Db 5676 AGGAACCTTCAAGAGGCGCACCGATGAGCTGCACTCAAGCTGCGCCCAAGCTGAGGTGATCA 5617
Qy 1242 AGGATCCTGGCAGCCCGTGGGCGATCTCTCTATGACTCTCTCCAAAGATCACCTCGAGA 1301
Db 5616 AGGATCCTGGCAGCCCGTGGGCGATCTCTCTATGACTCTCTCTGCAAGATCACCTTGAAA 5557
Qy 1302 AAGTCAAGGCACTTCGAGGAGAAATGGCGCTCTGAAAGAGACGCTGAGCAGCTCAATG 1361
Db 5556 AAGTCAAGGCACTTCGAGGAGAAATGGCGCTCTGAAAGAGATGCAATCTGCTGATG 5497
Qy 1362 ACCTTGCTCGCCAGCTTACCACTTTGGGATTCAGCTCTCACCGTATAACCTCAGCACTC 1421
Db 5496 ACCTTGCAATCATGCTGACCCACACTGGGCACTTCACTCTCACCTTATACCTCAGCACTT 5437


```
1422 TGAAGACCTGAAACACACAGATGAAGCTTTCTGAGGTGGCCGTGAGACCCAGTCAAGC 1481
Db
5436 TGAAGATCTGAATACAGATGAGGCTTCTACAGTGTGCTGGAGACCGTGTCAAGC 5377
2Y 1482 AGCTGATGAGCCCAACAGGACTTTGGTCCAGCATCTCAGACATTTCTTCCAGCTGT 1541
Db 5376 AGCTGATGAGCCCAACAGGACTTTGGTCCAGCATCTCAGACATTTCTTCCAGCTGT 5317
2Y 1542 TCCAGGTCCTTGGGAGAGAGCCATCTCCGCAAAACAAAGTGCCTTACTATATCAACACG 1601
Db 5316 TCCAGGTCCTTGGGAGAGAGCCATCTCACCACAAAGTGCCTTACTATATCAACACG 5257
2Y 1602 AGACTCAAAACATTTGCTGGGACCATCCCAAATGACAGATCTTACAGTCTTTAGTG 1661
Db 5256 AGACCCAAACCATTTGCTGGGACCATCCCAAATGACAGATCTTACAGTCTTTAGTG 5197
2Y 1662 ACCTGAATATGTCAGATTTCTCAGCTTTAGGACTGCCATGAATCTCCGAGACTGCAGA 1721
Db 5196 ACCTGAATATGTCAGTTCCTCGGTATAGGACTGCCATGAATCTCCGAGACTGCAGA 5137
2Y 1722 AGCCCTTTGCTGGATCTCTGAGCTCTCAGCTGCAATGTCATGTCCTTGGACCAACA 1781
Db 5136 AGCCCTTTGCTGGATCTCTGAGCTCTCAGCTGCAATGTCATGTCCTTGGACCAACA 5077
2Y 1782 ACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTTAATTTGTTGACCACTA 1841
Db 5076 ACCTCAAGCAAAATGACAGCCCATGATATCTCGAGATTTAATTTGTTGACCACTA 5017
2Y 1842 TTTATGACCCCTTGGAGAGAGCAACAAATTTGGTCAACCTCCTCTGCGTGGATA 1901
Db 5016 TTTATGATCTCTGGAGAGAGCAACAAATTTGGTCAACCTCCTCTGCGTGGATA 4957
2Y 1902 TGTGTCGAACCTGCTGCTGTAATTTATGATACGGGACGAACAGGAGGATTCGCTGTC 1961
Db 4956 TGTGTCGAACCTGCTGCTGTAATTTATGATACGGGACGAACAGGAGGATTCGCTGTC 4897
2Y 1962 TGTCTTTTAAACTGSCATCTTTCCCTGTAAGAGCAATTTGTAAGAGCAATTTGTAAG 2021
Db 4896 TGTCTTTTAAACTGSCATCTTTCCCTGTAAGAGCAATTTGTAAGAGCAATTTGTAAG 4837
2Y 2022 ACCTTTTCAAGCAAGTGGCAATTTCAACAGGATTTTGTGACCGCGCAGGTGGCCCTCC 2081
Db 4836 ACCTTTTCAAGCAAGTGGCAATTTCAACAGGATTTTGTGACCGCGCAGGTGGCCCTCC 4777
2Y 2082 TTTGTCATGATCTATCCAA 2101
Db 4776 TTTGTCATGATCTATCCAA 4757
```

```
RESULT 5
JS-09-091-501B-7
Sequence 7, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 6045
```

```
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (11)..(6037)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Chimeric
NAME/KEY: misc feature
LOCATION: (724)..(758)
OTHER INFORMATION: Precise residue is left open
US-09-091-501B-7
```

```
Query Match 26.1%; Score 547.4; DB 4; Length 6045;
Best Local Similarity 61.4%; Pred. No. 5.2e-159;
Matches 900; Conservative 0; Mismatches 556; Indels 9; Gaps 1;
QY 645 ACATGGTGGTGTAGTGTGATGTAATCTAGTGGAGATCAGCAACTGCTTGTGGAAGAAC 704
Db 3270 ACAGGCAAGAGATGGTAAAGCTTTGGGAAATTTCTGAAGAGGTACTATGCTTCAACATC 3329
QY 705 AACTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGG 764
Db 3330 GACTGGATGATATGAACCAAGATGGAATGACTTAAAGCAAAATCTGCTAGCATCAGGG 3389
QY 765 TTTCTTTTACAGACAGTTCTGACCACTGGAAGCGTGTGCACCTTTCTCTGACAGAACTTC 824
Db 3390 CCATTTGGAGGCGCAGCGCTGAGAAGTGGAAACAGGTTGCTGATGCTCTTAGAAGAACTGA 3449
QY 825 TGGTGTGGCTTACAGCTGAAAGATGATGAATTAAGCCGCGGAGGACCTATTGGAGGCGACT 884
Db 3450 TCAATGGCTGAATATGAAGATGAAGAGCTTAAAGCAAAATGCTTATGGAGGAGATG 3509
QY 885 TTCCAGAGTTCAGAGACAGAAACGATGTATAGGGCCCTTCAAGAGGGAATTTGAAACTA 944
Db 3510 TTCCAGCTTTACAGCTCCAGTATGACCATTTGAAGGCCCTTGAGACGGGAGTTAAAGGAGA 3569
QY 945 AAGAACTGTAAATCATGATGATCTTTGAGACTGTACGAAATATTTCTGACAGAGCAGCTT 1004
Db 3570 AAGAAATTTCTGCTGTAATGCTGTGCAAGCCGCGAGTTTCTTGGCTGATCAGCCAA 3629
QY 1005 T-----GGAAGACTAGAGAAACTCTACAGGAGCCCGAGAGAGCTGCTCTCTGAGG 1055
Db 3630 TTGAGGCCCTTGAAGAGCCAGAGAAACCTCAATCAAAACAGAAATTAACCTCTGAGG 3689
QY 1056 AGAGAGCCCAAGATGTCACTCGGCTTCTAGAAAGCAGCTGAGGAGGTCAATCTAGT 1115
Db 3690 AGAGAGCCCAAGATGTGCAAGCCATCGCAACAGCTCTTCTGAAAGTCAAAAGAAAAAT 3749
QY 1116 GGGAAAAATTGAACCTGCACTCGCTGACTGGGAGAGAAAAATAGATGAGACCTTTGAAA 1175
Db 3750 GGGAAAGTCTAAATGCTGTAACTAGCAATTTGGCAAAAGCAAGTGGNACAGGCATTGGAGA 3809
QY 1176 GACTCCAGGAATTTCAAGAGGCGACCGATGAGTGGAGCTCAAGTGCAGCCCAAGCTGAGG 1235
Db 3810 AACTCAGAGACCTTGCAAGGAGCTTATGGATGACCTGAGCGCTGACATGAAGAGGCGCAGT 3869
QY 1236 TGATCAAGGATCTCGCAGCCGCTGGGCGATCTCTCTCAATGACTCTCTCCAGATCAC 1295
Db 3870 CCGTGGGAATGCTGGAAGCCGCTGGGAGACTTACTCAATGACTCGCTGCGAGATCA 3929
QY 1296 TCAGAAAGTCAAGGCACTTCGAGGAGAAATTTGGCGCTCTGAAAGAGAACGTCGAGCCAG 1355
Db 3930 TTGAAAAAATCATGGCATTTAGAGAGAAATTTGCAACCAATCAACTTTAAAGTTAAACCG 3989
QY 1356 TCAATGACCTTGTCCGCGAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCA 1415
Db 3990 TGAATGATTTATCCAGTCAGCTGTCTCCACTTGAACCTGCTCTCTTAAAGATGTCTC 4049
QY 1416 GCACCTTGGAAAGACCTGAACACACAGATGGAAGCTTCTGAGGTGGCCGTCGAGGACCGAG 1475
Db 4050 GCCAGCTAGATGACCTTAATATCGATGGAACCTTTTACAGGTTTCTGTGGATGATCGCC 4109
```


1536 CGTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCA 1595
1537
8445 CGTCACTCCAGCTGCGTGGGCAAGATCCATTTACATATAAAGTGGCTTATATATCA 8504
1596 ACCAGAGACTCAAACTTCTCTGGGAGGACATCCCAAAATGACAGAGCTTACAGTCTTT 1655
1597
8505 ACCATCAAAACAGAGCCACTCTTGGGAGGACATCCCAAAATGACAGAGCTTACAGTCTTT 8564
1656 TAGCTGACCTGAATATGTCAGATCTCTAGCTTATAGGACTGCCATGAAATCTCCGAAGAC 1715
1657
8565 TTGCTGACCTGAATATGTCAGATCTCTAGCTTATAGGACTGCCATGAAATCTCCGAAGAC 8624
1716 TCGAGAGGCCCTTTGCTTGGATCTCTTGAAGCTCTGAGCTCTGATGATGATGATGATGATG 1775
1776
8625 TACAAAAGACATATGTTTGGATCTCTTAGAGTTGAGTACAAATGAAATTTTCAAC 8684
1777
1776 AGCAACACCTCAAGCAAAATGACAGCCATGATATCTCGAGATTTATTAATGTTTGA 1835
1777
8685 AGCAACAGTTGAACCAAAATGACAGCTCTCTAGTTTCCAGATGTCATCACTGTCTGA 8744
1836 CCATATTTATGACCCCTGGAGCAAGAGCAACAAATTTGGTCAAGCTCTCTCTGCG 1895
1837
8745 CAACAACATATGATGACCTTGAAGCAATGATGATGATGATGATGATGATGATGATGATG 8804
1896 TGGATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1955
1897
8805 TTGATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8864
1956 GTGCTCTCTTTTAAACTGTCATCAATTTCCCTGTGTAAGAGCAATTTGGAAGCAAGT 2015
1957
8865 GAGTGCAGAGTCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8924
2016 ACAGATACCTTTTCAAGCAAGTGGCAAGTTCAAGAGATTTTGTGACAGCGCAGGCTGG 2075
2076
8925 ACAGATATCTCTTTAAGCAAGTGGCGGCGGACAGCAAAATGTTGACAGAGGCGAGCTGG 8984
2077
2076 GCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
8985 GCCTGTACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9009

RESULT 7
US-09-091-501B-5
Sequence 5, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 200
TYPE: DNA
ORGANISM: Rattus sp.
US-09-091-501B-5
Query Match 3.8%; Score 79.4; DB 4; Length 200;
Best Local Similarity 64.3%; Pred. No. 1e-14;
Matches 119; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
565 CCTAAAGCCCAAGTACACACATAGAGTCTTCAAGAGATCTAGAACAGACAAAGT 624

Db 16 CCTGCAAAACCTCTTGAAGAACATATAAGTTTGCAAAAGTGAACCTGGAAGCTGAGCAGGT 75
QY 625 CAGGGTCAATTTCTCTCACTACATGATGATGATGATGATGATGATGATGATGATGATGATG 684
Db 76 GAAGGTGAATTTCTTAACTCATATGATGATGATGATGATGATGATGATGATGATGATGATG 135
QY 685 AACTGTGCTTTTGAAGAACAACTTAAGCTATTGGGAGATCGATGGGCAAAACATCTGTAG 744
Db 136 CACAGCTGTTTGAAGATCAGTTACAGAACTGGGTGAGCGCTGGACAGCTGTATGCCG 195
QY 745 ATGGA 749
Db 196 CTGGA 200
RESULT 8
US-09-091-501B-4
Sequence 4, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrrophin gene expression
FILE REFERENCE: 620-42
CURRENT APPLICATION NUMBER: US/09/091,501B
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/GB96/03156
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: GB 9525962.8
PRIOR FILING DATE: 1995-12-19
PRIOR APPLICATION NUMBER: GB 9615797.9
PRIOR FILING DATE: 1996-07-26
PRIOR APPLICATION NUMBER: GB 9622174.2
PRIOR FILING DATE: 1996-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 200
TYPE: DNA
ORGANISM: Mus sp.
US-09-091-501B-4
Query Match 3.7%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.8e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 553 TGATCTTGAAGACCTAAACGCCAAGTACAAACATAGAGTCTTCAAGAGATCTAGA 612
Db 4 TGACCTGCCCTCCCTGCAGAGCTGTTCAAGAACATAAAAGTTTGCAAAATGACCTTGA 63
QY 613 ACAAGAACAGTCAAGGTCAATTTCTCACTCAGATGATGATGATGATGATGATGATGATG 672
Db 64 AGCTGAACAGGTGAAGGTAAATTTCTTAACTCAGATGATGATGATGATGATGATGATGATG 123
QY 673 TGGAGATCAGCAACTGCTGCTTTTGAAGAACAACTTAAGGTATTGGGAGATCGATGGGC 732
Db 124 TGGGAGAGTGCACAGCTCTTCTGGAAGATCAGTTACAGAACTGGGTGAGCGCTGGAC 183
QY 733 AAACATCTGTAGATGA 749
Db 184 AGCTGTATCCGCTGA 200
RESULT 9
US-09-091-501B-6
Sequence 6, Application US/09091501B
Patent No. 6518413
GENERAL INFORMATION:
APPLICANT: Tinsley, Jonathon M
APPLICANT: Davies, Kay E
TITLE OF INVENTION: Utrrophin gene expression
FILE REFERENCE: 620-42

471 ATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAAAGAAACAAGGAAA 530

Db 602 AGCAAGATATAAGATTTGATCGATCAGACAAAGAAAAATGGAGATACGATCGAGGAA 661
2Y 531 TGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACTAAACCGCAAGGTACACACATA 590
Db 662 TTGTAGAAGTGTCTGTTGAGGGCTTCAGCTGGATTGGAAGCTTACGTACAAATGGGACA 721
2Y 591 AGTGCTTCAAGAGATCTAGAACCAAGAACAAAGTCAAGGTCAATTTCTCTCACTCACTGG 650
Db 722 CGAAGCTAGATGCCAAATCGCACAGCTGTGTAGTATCAATGCTTTAAAGCGGTAG 781
2Y 651 TGTGTGTAGTTGATGAATCTAGTGG 675
Db 782 AATTGGGTTCGGATTCACCTTTCTGG 806

RESULT 14
US-09-620-312D-69
; Sequence 69, Application US/09620312D
; Patent No. 6589662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6589662el Nucleic Acids and
Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 69
LENGTH: 1690
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (128)..(1522)

Query Match 2.0%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.0098;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
Y 1180 CAGGAATTTCAAGAGGCCACCGATGAGCTGAGCTCAAGCTGCGCCAAAGCTGAGGTGAT 1239
b 55 CAAGGAGTTGACCAAGTGGCGCAGCACTGGACGACGAGCTGCGATGGGTTCAGGAGCG 114
Y 1240 CAAGGGA---TCCTGGAGCCCGTGGGGATCTCTCAATTGACTCTCTCCAGATCACT 1296
b 115 GCTGCCACTGGCCATGCAGACAGACCGGAGCAACCGTTTGCAGCGGTCCAGCAGCACT 174
Y 1297 CGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACTGAGCCACGT 1356
b 175 CAAAAGAACCCAGGGCCTTCGCGGGGAGATCCAGGCGCATGGGCGCGCTTGGAGGAGT 234

QY 1357 CAATGACCTTGTCTGCGCAGCTTACCACTTTGGCACTTCACTCTCACCGTATAACCTCAG 1416
Db 235 GCTGGAGCGCGCGGGCGCTGCGCTGCGCTGCGAGCCCGGAGGAGAGGCACTGCGCG 294
QY 1417 CACTCTGGAAGACCTGAACACCCAGATGGAACTTCTGCAGGTGGCGCTCGAGGACCGAGT 1476
Db 295 GGGCCTGGAGCAGCTGCAGAGCGCTGGGCGGCACTGCGGAGGCTGCCGAGCGGCA 354
QY 1477 CAGGCACTGCATGAAGCCACAGGG 1502
Db 355 GCAGGTGCTGGACGCCCGCTTCCAGG 380

RESULT 15
US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(102)
; NAME/KEY: unsure
; LOCATION: (1021)...(1023)
; NAME/KEY: unsure
; LOCATION: (2266)...(2268)
US-09-368-590-1

Query Match 2.0%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.027;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
QY 1180 CCAGGAATTTCAAGAGGCCACCGATGAGCTGAGCTCAAGCTGCGCCAAAGCTGAGGTGAT 1239
Db 3639 CAAGGAGTTGCACCAAGTGGCGCACGACTGGACGAGCTGGATGGGTTCAGGAGCG 3698
QY 1240 CAAGGGA---TCCTGGAGCCCGTGGGGATCTCTCAATTGACTCTCTCAAGATCACT 1296
Db 3699 GCTGCCACTGGCCATGCAGACAGACGAGCGCAACGGTTTGCAGGGCGGTCCAGCAGCAT 3758
QY 1297 CGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGCTCTGAAAGAGAACTGAGCCACT 1356
Db 3759 CAAAAGAACCCAGGGCTTCGCGCGGAGATCCAGCGCATGGGCGCGCTTGGAGGAGT 3818
QY 1357 CAATGACCTTGTCTGCCAGCTTACCACTTTTGGGCACTCAGCTTCACCGTATAACCTCAG 1416
Db 3819 GCTGAGCGCGCGGGCGGCTGCGCTGCTGCGAGCCCGGAGGAGAGGAGTTCGCGCG 3878
QY 1417 CACTCTGGAAGACCTGAACACCCAGATGGAACTTCTGCAGGTGGCGCTCGAGGACCGAGT 1476
Db 3879 GGGCTTGGAGAGCTGCAGAGCGCTGGCGCGGACTGCGGGAGGCTGCCGAGCGGCA 3938
QY 1477 CAGGCACTGCATGAAGCCACAGGG 1502
Db 3939 GCAGGTGCTGAGCGCGCTTCCAGG 3964

Search completed: February 2, 2004, 06:16:01
Job time : 133.879 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 23:51:42 ; Search time 687.863 Seconds
(without alignments)
11131.886 Million cell updates/sec

Title: US-09-845-416-12_COPY_900_3000

Perfect score: 2101
Sequence: 1 gagctatgctacacagg.....ttctgcatgtatccaa 2101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2101	100.0	3510	13	US-09-845-416-12
2	2101	100.0	4476	13	US-09-845-416-31
3	1760.4	83.8	5339	13	US-10-149-736-40
4	1750.8	83.3	3531	13	US-09-845-416-10
5	1750.8	83.3	4498	13	US-09-845-416-30
6	1743	83.0	3858	13	US-09-845-416-9
7	1743	83.0	4825	13	US-09-845-416-29
8	1743	83.0	4848	13	US-09-845-416-35
9	1743	83.0	5060	13	US-09-845-416-36
10	1641	78.1	4414	13	US-09-845-416-32
11	1629	77.5	5462	13	US-09-845-416-14
12	1629	77.5	5462	13	US-10-149-736-41
13	1611.6	76.7	5417	13	US-10-149-736-39
14	1602	76.2	3959	13	US-09-845-416-6
15	1602	76.2	4966	13	US-09-845-416-28

16	1602	76.2	4990	13	US-09-845-416-34	Sequence 34, Appl
17	1419	67.5	4182	13	US-09-845-416-2	Sequence 2, Appl
18	1419	67.5	5149	13	US-09-845-416-27	Sequence 27, Appl
19	1328.8	63.2	2169	13	US-09-845-416-4	Sequence 4, Appl
20	1328.8	63.2	8689	13	US-10-149-736-42	Sequence 42, Appl
21	1328.8	63.2	11058	13	US-09-845-416-1	Sequence 1, Appl
22	1328.8	63.2	11443	13	US-10-149-736-44	Sequence 44, Appl
23	1328.8	63.2	12057	13	US-10-149-736-47	Sequence 47, Appl
24	1328.8	63.2	13957	10	US-09-782-378A-22	Sequence 22, Appl
25	1328.8	63.2	13957	10	US-09-880-107-2284	Sequence 2284, Ap
26	1328.8	63.2	13957	13	US-10-149-736-1	Sequence 1, Appl
27	1328.8	63.2	14082	13	US-10-341-434-108	Sequence 108, App
28	1324	63.0	1821	13	US-09-845-416-13	Sequence 13, Appl
29	1149.6	54.7	13815	13	US-10-149-736-2	Sequence 2, Appl
30	937	44.6	1434	13	US-09-845-416-15	Sequence 15, Appl
31	785.4	37.4	1991	13	US-09-845-416-3	Sequence 3, Appl
32	777	37.0	1667	13	US-09-845-416-7	Sequence 7, Appl
33	554.2	26.4	11096	13	US-10-149-736-4	Sequence 4, Appl
34	547.4	26.1	10302	10	US-09-782-378A-23	Sequence 23, Appl
35	547.4	26.1	10302	13	US-10-149-736-3	Sequence 3, Appl
36	538.6	25.6	16531	13	US-10-101-510-667	Sequence 667, App
37	450	21.4	1340	13	US-09-845-416-11	Sequence 11, Appl
38	393	18.7	887	13	US-10-149-736-35	Sequence 35, Appl
39	387	18.4	387	13	US-10-149-736-32	Sequence 32, Appl
40	331	15.8	333	13	US-10-149-736-9	Sequence 9, Appl
41	327	15.6	327	13	US-10-149-736-8	Sequence 8, Appl
42	324	15.4	324	13	US-10-149-736-33	Sequence 33, Appl
43	216	10.3	216	13	US-10-149-736-34	Sequence 34, Appl
44	190.2	9.1	256	9	US-09-864-761-21956	Sequence 21956, A
45	178	8.5	456	9	US-09-864-761-6092	Sequence 6092, Ap

ALIGNMENTS

RESULT 1

US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845, 416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200, 777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match	100.0%;	Score 2101;	DB 13;	Length 3510;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2101;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GAGCTATGCTACACACAGCTGCTTATGTGTCACACCTTGACCTTACACGAGCCCAIT	60	
Db	900	GAGCTATGCTTACACACAGCTGCTTATGTGTCACACCTTGACCTTACACGAGCCCAIT	959	
Qy	61	TCCTTACACACATTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTTCATTGATGAGAG	120	
Db	960	TCCTTACACACATTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTTCATTGATGAGAG	1019	
Qy	121	TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGTATTATCGTGGCTTTCTTC	180	
Db	1020	TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAAGTATTATCGTGGCTTTCTTC	1079	
Qy	181	TGCTGAGGACACATTGCAAGCAGAGAGATTTCTAATGATGTCGAAGTGTGTAAGA	240	

1080 TGCTGGGACACATTGCAAGCAGGAGAGATTCTTAATGATGTGGAGTGTGGAAGA 1139
241 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACACCCCATCAGGGCCGGGTGG 300
1140 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACACCCCATCAGGGCCGGGTGG 1199
301 TAAATATTCTACAAATGGGAAGTAAGCTGATTTGGAACAGAGAAAATTATCAGAAAGATGAAGA 360
1200 TAAATATTCTACAAATGGGAAGTAAGCTGATTTGGAACAGAGAAAATTATCAGAAAGATGAAGA 1259
361 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGAGTGGGAATGCCCTCAGGGTAGC 420
1260 AACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGAGTGGGAATGCCCTCAGGGTAGC 1319
421 TAGCATGGAAGAAAACAAAGCAATTTACATAGAGTGTAAATGATCTCCAGAAATCAGAAACT 480
1320 TAGCATGGAAGAAAACAAAGCAATTTACATAGAGTGTAAATGATCTCCAGAAATCAGAAACT 1379
481 GAAGAGTTGAATGACTGTGCTAACAAAACAGAGAAAGAAACAAGGAAATGGAGGAAGA 540
1380 GAAGAGTTGAATGACTGTGCTAACAAAACAGAGAAAGAAACAAGGAAATGGAGGAAGA 1439
541 GCCTCTTGACCTGATCTTGAGACCTAAACGCCCAAGTACAGCAACATAGAGTGTCTTCA 600
1440 GCCTCTTGACCTGATCTTGAGACCTAAACGCCCAAGTACAGCAACATAGAGTGTCTTCA 1499
601 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 660
1500 AGAAGATCTAGAACAAAGAACAGTCAAGGTCAATTTCTCTCACTCAGATGGTGGTAGT 1559
661 TGATGAATCTAGTGAAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAGAGTATTGGG 720
1560 TGATGAATCTAGTGAAGATCAGCAACTGCTGCTTTGGAAGAACAACTTAGAGTATTGGG 1619
721 AGATCGATGGGCAAAACATCTGTAGATGACAGAACACCGCTGGGTCTTTTACAGACAG 780
1620 AGATCGATGGGCAAAACATCTGTAGATGACAGAACACCGCTGGGTCTTTTACAGACAG 1679
781 TTCTGACAGTGGAGCGTCTGCACTTTCTTGAGGAACCTCTGAGTGGGTGAGTACAGCT 840
1680 TTCTGACAGTGGAGCGTCTGCACTTTCTTGAGGAACCTCTGAGTGGGTGAGTACAGCT 1739
841 GAAAGATGATGAATTAAGCCGCGAGCCACTTATGGAGGCACTTTCCAGCAGTTCAGAA 900
1740 GAAAGATGATGAATTAAGCCGCGAGCCACTTATGGAGGCACTTTCCAGCAGTTCAGAA 1799
901 GCAGAACGATGTACATAGGGCCTTCAAGAGGAATGAAAACCTAAAGAACCTGTAATCAT 960
1800 GCAGAACGATGTACATAGGGCCTTCAAGAGGAATGAAAACCTAAAGAACCTGTAATCAT 1859
961 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGCAGCCTTTGGAAGGACTTAGAGAA 1020
1860 GAGTACTCTTGAGACTGTACGAATATTTCTGACAGCAGCCTTTGGAAGGACTTAGAGAA 1919
1021 ACTCTACAGGAGCCAGAGAGCTCCTCCTGAGGAGAGCCAGAAATGTCACTCGGCT 1080
1920 ACTCTACAGGAGCCAGAGAGCTCCTCCTGAGGAGAGCCAGAAATGTCACTCGGCT 1979
1081 TCTACGAAGACGGCTGAGGAGGTCAATCTAGTGGGAAAATTTGAACCTGCATCCGC 1140
1980 TCTACGAAGACGGCTGAGGAGGTCAATCTAGTGGGAAAATTTGAACCTGCATCCGC 2039
1141 TGACTGGCAGAGAAAATAGATGAGACCTTGAAGACTCCAGGAACCTCAAGAGGCCAC 1200
2040 TGACTGGCAGAGAAAATAGATGAGACCTTGAAGACTCCAGGAACCTCAAGAGGCCAC 2099
1201 GATGAGCTGGACCTCAAGCTGGCCAGCTGAGTGTATCAAGGGATCCTGGCAGCCCGT 1260
2100 GATGAGCTGGACCTCAAGCTGGCCAGCTGAGTGTATCAAGGGATCCTGGCAGCCCGT 2159
1261 GGGCGATCTCCTCATTTGACTCTCTCAAGATCAGCTCGAGAAAGTCAAGGGACATTCGAGG 1320

2160 GGGCGATCTCCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGCAGCATTCGAGG 2219
1321 AGAAATTTGGCCCTCTGAAAGAGAAAGTGGAGCCACCTCAATGACCTTGTCTGCGACCTTAC 1380
2220 AGAAATTTGGCCCTCTGAAAGAGAAAGTGGAGCCACCTCAATGACCTTGTCTGCGACCTTAC 2279
1381 CACTTTTGGSCATTTCAGCTCTCAACCGTATAAACCCTCAGCACTCTGGAAGACCTTGAACACCAG 1440
2280 CACTTTTGGSCATTTCAGCTCTCAACCGTATAAACCCTCAGCACTCTGGAAGACCTTGAACACCAG 2339
1441 ATGGAGCTTCTGAGAGTGGCCGCTCGAGGACCGAGTCAAGGAGCTGCATGAGCCACAG 1500
2340 ATGGAGCTTCTGAGAGTGGCCGCTCGAGGACCGAGTCAAGGAGCTGCATGAGCCACAG 2399
1501 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAG 1560
2400 GGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGCTCTGTCCAGGGTCCCTGGGAGAG 2459
1561 AGCCATCTCGCCAAAACAAAGTCCCTACTATATCAACCACGAGACTCAACAACTTGTCTG 1620
2460 AGCCATCTCGCCAAAACAAAGTCCCTACTATATCAACCACGAGACTCAACAACTTGTCTG 2519
1621 GGACCATCCCAAAATCACAAGACTCTACAGTCTTTAGCTGACCTGAATATGTCCAGATT 1680
2520 GGACCATCCCAAAATCACAAGACTCTACAGTCTTTAGCTGACCTGAATATGTCCAGATT 2579
1681 CTGAGCTTATAGGACTGCGCATGAAACTCCGAAAGATGCGAAGAGCCCTTTGGTTCGATCT 1740
2580 CTGAGCTTATAGGACTGCGCATGAAACTCCGAAAGATGCGAAGAGCCCTTTGGTTCGATCT 2639
1741 CTGAGCCTGTGAGCTGCGATGATGCTTGGACGAGCAACACCTCAAGCAAAATGACCA 1800
2640 CTGAGCCTGTGAGCTGCGATGATGCTTGGACGAGCAACACCTCAAGCAAAATGACCA 2699
1801 GCCCATGATATCTCTGAGATTATTAATTTGTGACCACTATTTATGACCGCTCGAGCA 1860
2700 GCCCATGATATCTCTGAGATTATTAATTTGTGACCACTATTTATGACCGCTCGAGCA 2759
1861 AGAGCACACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATATGCTGAACTGGCTGCT 1920
2760 AGAGCACACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATATGCTGAACTGGCTGCT 2819
1921 GAATGTTATGATAGCGGACGAAACAGGAGGATCCGTGTCTCTGCTTTTAAACTGGCAT 1980
2820 GAATGTTATGATAGCGGACGAAACAGGAGGATCCGTGTCTCTGCTTTTAAACTGGCAT 2879
1981 CATTTCCCTGTGTAAGCAGCATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
2880 CATTTCCCTGTGTAAGCAGCATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC 2939
2041 AAGTTCAACAGGATTTTGTGACGAGCGAGGCTGGGCTCCTTCTGCGATGATTCTATCCA 2100
2940 AAGTTCAACAGGATTTTGTGACGAGCGAGGCTGGGCTCCTTCTGCGATGATTCTATCCA 2999
2101 A 2101
3000 A 3000

RESULT 2

US-09-845-416-31
; Sequence 31, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR APPLICATION NUMBER: 2001-04-30
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 31

LENGTH: 4476

TYPE: DNA

ORGANISM: Homo sapiens

IS-09-845-416-31

Query Match

Best Local Similarity 100.0%; Score 2101; DB 13; Length 4476;

Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1	GAGCTATGCTTACACAGAGCTGCTTATGTATGATCAGCACTCTGACCCCTACAGGAGCCCATTT	60
1656	GAGCTATGCTTACACAGAGCTGCTTATGTATGATCAGCACTCTGACCCCTACAGGAGCCCATTT	1715
61	TCCTTTCACAGCATTTTGAAGCTCCCTGAAGACAAGTCATTTGGCAGTTCATTTGATGGAGAG	120
1716	TCCTTTCACAGCATTTTGAAGCTCCCTGAAGACAAGTCATTTGGCAGTTCATTTGATGGAGAG	1775
121	TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAAGTATATCGTGGCTTCCTTC	180
1776	TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAGAAGTATATCGTGGCTTCCTTC	1835
181	TGCTGAGGACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA	240
1836	TGCTGAGGACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA	1895
241	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGGTTGG	300
1896	CCAGTTTTCATCTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGCGGGTTGG	1955
301	TAAATTTCTACAAATTTGGAGTAAGCTGATTTGGAACAGAGAAATTTATCAGAAGATGAAGA	360
1956	TAAATTTCTACAAATTTGGAGTAAGCTGATTTGGAACAGAGAAATTTATCAGAAGATGAAGA	2015
361	AACTGAAAGTACAAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCCTCAGGGTAGC	420
2016	AACTGAAAGTACAAGAGCAGATGAATCTCTCTAAATTCAGATGGGAATGCCCTCAGGGTAGC	2075
421	TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATACAGAACT	480
2076	TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATACAGAACT	2135
481	GAAGAGTTGATGATCTGCTTAAACAAACAGAGAGAGAAACAGAGAAATTTGAGGAGGA	540
2136	GAAGAGTTGATGATCTGCTTAAACAAACAGAGAGAGAAACAGAGAAATTTGAGGAGGA	2195
541	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCGCAAGTACAAACATAGAGTGTCTCA	600
2196	GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCGCAAGTACAAACATAGAGTGTCTCA	2255
601	AGAAGATCTAGAACAGAAACAAAGTCAAGTCAAGTCAATCTCTCACTCAGTGTGGTGTAGT	660
2256	AGAAGATCTAGAACAGAAACAAAGTCAAGTCAAGTCAATCTCTCACTCAGTGTGGTGTAGT	2315
661	TGATGAATCTAGTGAGATCAAGCACTGCTTGTGGAAGAAACAACTTAAGGTATTGGG	720
2316	TGATGAATCTAGTGAGATCAAGCACTGCTTGTGGAAGAAACAACTTAAGGTATTGGG	2375
721	AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCCTTTTACAGACAG	780
2376	AGATCGATGGGCAACATCTGTAGATGACAGAGACCGCTGGGTTCCTTTTACAGACAG	2435
781	TTCTGACAGTGAAGCGTCTGCACTTCTCTGAGGAACTCTCTGGTGTGGCTTACAGCT	840
2436	TTCTGACAGTGAAGCGTCTGCACTTCTCTGAGGAACTCTCTGGTGTGGCTTACAGCT	2495
841	GAAGATGATGAATTTAAGCGGCAAGCACTTATGGAGCGCACTTTCCAGCAGTTTCAGAA	900
2496	GAAGATGATGAATTTAAGCGGCAAGCACTTATGGAGCGCACTTTCCAGCAGTTTCAGAA	2555
901	CGAGAACGATGTACATAGGCTTCAAGAGGAAATTTGAAGAACTTAAGAACTGTATCAT	960

Db	2556	GCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACTTAAAGAACCTGTATCAT	2615
QY	961	GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGAGCGCTTTGGAAGGACTAGAGAA	1020
Db	2616	GAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGAGCGCTTTGGAAGGACTAGAGAA	2675
QY	1021	ACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCT	1080
Db	2676	ACTCTACAGGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCT	2735
QY	1081	TCTACAAAGAGCGCTGAGAGAGTCAATACTGAGTGGGAAAAATTTGAACTCGCACTCCGC	1140
Db	2736	TCTACAAAGAGCGCTGAGAGAGTCAATACTGAGTGGGAAAAATTTGAACTCGCACTCCGC	2795
QY	1141	TGACTGGCAGAGAAAAATAGATGAGACCTTCAAAAGACTTCCAGGAACTTCAAGAGGCCAC	1200
Db	2796	TGACTGGCAGAGAAAAATAGATGAGACCTTCAAAAGACTTCCAGGAACTTCAAGAGGCCAC	2855
QY	1201	GGATGAGCTTGGACCTCAAGCTGGCGCAAGCTGAGGTGATCAAGGGATCTGGAGGCCGT	1260
Db	2856	GGATGAGCTTGGACCTCAAGCTGGCGCAAGCTGAGGTGATCAAGGGATCTGGAGGCCGT	2915
QY	1261	GGCGGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG	1320
Db	2916	GGCGGATCTCTCATTTGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG	2975
QY	1321	AGAAATTTGGCGCTCTGAAAGAGACGTGAGCCACGTCAATGACCTTGTCTGCGCAGCTTAC	1380
Db	2976	AGAAATTTGGCGCTCTGAAAGAGACGTGAGCCACGTCAATGACCTTGTCTGCGCAGCTTAC	3035
QY	1381	CAGTTTGGCGATTCAGCTCTCAAGTATTAACCTCAGCACTCTGGAAGACCTGAACACACAG	1440
Db	3036	CAGTTTGGCGATTCAGCTCTCAAGTATTAACCTCAGCACTCTGGAAGACCTGAACACACAG	3095
QY	1441	ATGGAAGCTTTCGAGGTGGCGCTGAGACCGAGTCAAGGCACTGCAATGAAGCCACACAG	1500
Db	3096	ATGGAAGCTTTCGAGGTGGCGCTGAGACCGAGTCAAGGCACTGCAATGAAGCCACACAG	3155
QY	1501	GGACTTTGGTCAGCATCTCAGCAGCTTTCTTCCAGCTGTCTCAGGGTCCCTGGGAGAG	1560
Db	3156	GGACTTTGGTCAGCATCTCAGCAGCTTTCTTCCAGCTGTCTCAGGGTCCCTGGGAGAG	3215
QY	1561	AGCCATCTCGCAAAACAAAGTGCCTTACTATATCAACACAGAGCTCAAAACAACTTGTG	1620
Db	3216	AGCCATCTCGCAAAACAAAGTGCCTTACTATATCAACACAGAGCTCAAAACAACTTGTG	3275
QY	1621	GGACCAATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGGAATTAATGTGAT	1680
Db	3276	GGACCAATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGGAATTAATGTGAT	3335
QY	1681	CTCAGCTTATAGAGCTGCCATGAAACTCCGAGAGCTGAGAGGCGCTTTGCTTGGATCT	1740
Db	3336	CTCAGCTTATAGAGCTGCCATGAAACTCCGAGAGCTGAGAGGCGCTTTGCTTGGATCT	3395
QY	1741	CTTGAGCTGTGAGCTGATGTGATGCTTGGACAGCAGCAACCTCAAGCAAAATGACCA	1800
Db	3396	CTTGAGCTGTGAGCTGATGTGATGCTTGGACAGCAGCAACCTCAAGCAAAATGACCA	3455
QY	1801	GCCATGATATCTCGAGATTAATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA	1860
Db	3456	GCCATGATATCTCGAGATTAATTAATTTGTTGACCACTATTTATGACCGCTGGAGCA	3515
QY	1861	AGAGCAACAATTTTGGTCAACCGCTCTCTGCGTGGATATGTCTGAACTGGCTGCT	1920
Db	3516	AGAGCAACAATTTTGGTCAACCGCTCTCTGCGTGGATATGTCTGAACTGGCTGCT	3575
QY	1921	GAATGTTTATGATACGGAGCGAAACAGGAGAGATCCGTGCTCTGCTTTTAAACCTGGAT	1980
Db	3576	GAATGTTTATGATACGGAGCGAAACAGGAGAGATCCGTGCTCTGCTTTTAAACCTGGAT	3635
QY	1981	CATTTCCCTGTGTAAGGACATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC	2040
Db	3636	CATTTCCCTGTGTAAGGACATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGC	3695

QY 2041 AGTTCAACAGGATTTCTGACACGCGAGGCTGGGCTCTTCTGTCATGATTCATCCA 2100
DB 3696 AAGTTCAACAGGATTTCTGACACGCGAGGCTGGGCTCTTCTGTCATGATTCATCCA 3755
QY 2101 A 2101
DB 3756 A 3756

RESULT 3
US-10-149-736-40
; Sequence 40, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 5339
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-40

Query Match 83.8%; Score 1760.4; DB 13; Length 5339;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 191; Indels 15; Gaps 2;

QY 1 GAGCTATGCTACACACAGGCTGCTTATGTACCACCTCTGACCTACACGAGGCCCAT 60
DB 1099 GAGCTATGCTACACACAGGCTGCTTATGTACCACCTCTGACCTACACGAGGCCCAT 1158
QY 61 TCCTTCACAGCATTTGGAGCTCTGAAGCAAGTCATTTGGCAGTTCATTTGATGGAGAG 120
DB 1159 TCCTTCACAGCATTTGGAGCTCTGAAGCAAGTCATTTGGCAGTTCATTTGATGGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAGTATTTATCGTGGCTTCTTTC 180
DB 1219 TGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAAGAGTATTTATCGTGGCTTCTTTC 1278
QY 181 TGCTGAGGACATTTGCAAGCAAGGAGAGATTTCTAATGATGGAGTGGTGAAGA 240
DB 1279 TGCTGAGGACATTTGCAAGCAAGGAGAGATTTCTAATGATGGAGTGGTGAAGA 1338
QY 241 CCAGTTTCTATCTCATGAGGCTGATGATGATTTGACCCATCAGGCGCGGTTGG 300
DB 1339 CCAGTTTCTATCTCATGAGGCTGATGATGATTTGACCCATCAGGCGCGGTTGG 1398
QY 301 TAAATTTCTAATTTGGGAAGTAACTGATTTGAACAGGAGAAATTTATCAGAAGTGAAGA 360
DB 1399 TAAATTTCTAATTTGGGAAGTAACTGATTTGAACAGGAGAAATTTATCAGAAGTGAAGA 1458
QY 361 AACTGAGTACAGAGCAGATGAATCTCTAATTTCAAGATGGGAATGCCCTCAGGGTAGC 420
DB 1459 AACTGAGTACAGAGCAGATGAATCTCTAATTTCAAGATGGGAATGCCCTCAGGGTAGC 1518
QY 421 TAGCATGGAAGAAACAAAGCAATTTATCATAGATTTTAAATGGATCTCCAGAAATCAGAACT 480
DB 1519 TAGCATGGAAGAAACAAAGCAATTTATCATATAGATTTTAAATGGATCTCCAGAAATCAGAACT 1578
QY 481 GAAAGATTTGAATGATGGCTAAACAAACAGAGAAAGAAACAGGAAATGGAGGAAGA 540

DB 1579 GGAAAAATTTCTTGCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCTCTACAGGATGC 1638
QY 541 GCCTCTTGGACCTGATCTTTGAAGACCTAAACAGCCCACTAACAACATAGGTCCTTCA 600
DB 1639 TACCCGTGAAGAAAGGCTCCTAGAAGACTCCAAGGAGTAAAGAGCTGATGAACAATG 1698
QY 601 AGAAGATCTAGAACCAAGTCAAGTCAAGGCTCAATTTCTCTCACTCACTCACTGGTGGTAGT 660
DB 1699 GCAAGACCTCCAAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAA 1758
QY 661 TGATGA-----ATCTAGTGGAGATCAAGCAACTGCTGCTTTGGAGACAACT 708
DB 1759 CAGCCAAAAAATCTGAGATCCCTGGNAGGTTCCGATGATGAGCTCTGTACAAAGACG 1818
QY 709 TAAGG---TAITGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAAAGCTCTCTCAACATFAGTGC 765
DB 1819 TTTGATACATGAATCTCAAGTGAAGTGAATTCGGAAAAAGTCTCTCAACATFAGTGC 1878
QY 766 TTTTTTACAAGACAGTTCTGACCAAGTGAAGGCTGCACTTTCTCTGCAAGAACTTCT 825
DB 1879 CCATTTGGAAGCCAGTTCTGACCAAGTGAAGGCTGCACTTTCTCTGCAAGAACTTCT 1938
QY 826 GGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGGCACTTATGGAGGCGACTT 885
DB 1939 GGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGGCACTTATGGAGGCGACTT 1998
QY 886 TCCAGCAGTTTCAAGACGAAACGATGTATATAGGGCTTCAAGAGGGAATTTGAAACCTAA 945
DB 1999 TCCAGCAGTTTCAAGACGAAACGATGTATATAGGGCTTCAAGAGGGAATTTGAAACCTAA 2058
QY 946 AGAAGCTGTAATCATGAGTACTTCTGAGACTGTAGCAATATTTCTGACAGAGCAGCTTT 1005
DB 2059 AGAAGCTGTAATCATGAGTACTTCTGAGACTGTAGCAATATTTCTGACAGAGCAGCTTT 2118
QY 1006 GGAAGGACTTAGAAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCA 1065
DB 2119 GGAAGGACTTAGAAGAACTCTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCA 2178
QY 1066 GAATGTCACTCGGCTTCTACGAAAGGCTGAGAGTCAATCTGAGTGGGAAAAATTT 1125
DB 2179 GAATGTCACTCGGCTTCTACGAAAGGCTGAGAGTCAATCTGAGTGGGAAAAATTT 2238
QY 1126 GAACCTGCACTCGGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGA 1185
DB 2239 GAACCTGCACTCGGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGA 2298
QY 1186 ACTTCAAGAGGCGCAGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGTGAAGGG 1245
DB 2299 ACTTCAAGAGGCGCAGGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGTGAAGGG 2358
QY 1246 ATCTTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAGAGATCACCTCGAGAAAGT 1305
DB 2359 ATCTTGGCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAGAGATCACCTCGAGAAAGT 2418
QY 1306 CAAGGCACCTTCGAGGAGAAATTTGCGCTCTGAAAGAGAACTGAGCCACGTCATGACCT 1365
DB 2419 CAAGGCACCTTCGAGGAGAAATTTGCGCTCTGAAAGAGAACTGAGCCACGTCATGACCT 2478
QY 1366 TGCTGCGCAGCTTACACTTTGGGCAATTCAGTCTCAAGTATACCTCAGGACTCTGGA 1425
DB 2479 TGCTGCGCAGCTTACACTTTGGGCAATTCAGTCTCAAGTATACCTCAGGACTCTGGA 2538
QY 1426 AGACCTTGAACACAGATGGAAGCTTTCTGAGGTGGCCCTCGAGGAGCCGAGTCAGGAGCT 1485
DB 2539 AGACCTTGAACACAGATGGAAGCTTTCTGAGGTGGCCCTCGAGGAGCCGAGTCAGGAGCT 2598
QY 1486 GCATGAAGCCACAGGACTTTGGTTCAGCATCTCAGCACTTTCTTTTCAAGTCTGTCCA 1545
DB 2599 GCATGAAGCCACAGGACTTTGGTTCAGCATCTCAGCACTTTCTTTTCAAGTCTGTCCA 2658
QY 1546 GGTGCTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACACAGAGAC 1605
DB 2659 GGTGCTCCCTGGGAGAGAGCCATCTCGCCAAAACAAAGTGCCTACTATATCAACACAGAGAC 2718

1606 TCAAAACAATTCCTGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCT 1665
2719 TCAAAACAATTCCTGGAGCCATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCT 2778
1666 GAATAATGTCAGATTCCTAGCTTATAGGATCCCATGAATCCGAAGAGCTCAGAGGC 1725
2779 GAATAATGTCAGATTCCTAGCTTATAGGATCCCATGAATCCGAAGAGCTCAGAGGC 2838
1726 CTTTCTCTGGATCTCTTGGAGCTGTGAGCTGCATGTGATGCTTTGGACACAGCAACCT 1785
2839 CTTTCTCTGGATCTCTTGGAGCTGTGAGCTGCATGTGATGCTTTGGACACAGCAACCT 2898
1786 CRAGCAAAATGACACGCCCATGGATATCTCTGAGATTAATTAATTTGTCACCACTATTTA 1845
2899 CRAGCAAAATGACACGCCCATGGATATCTCTGAGATTAATTAATTTGTCACCACTATTTA 2958
1846 TGACCCCTCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCTGGATATGTG 1905
2959 TGACCCCTCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCTGGATATGTG 3018
1906 TCTGAATCTGGCTGCTGAATGTTTATGATACGGAGCAACAGGAGGATCCGTGCTCTGTC 1965
3019 TCTGAATCTGGCTGCTGAATGTTTATGATACGGAGCAACAGGAGGATCCGTGCTCTGTC 3078
1966 TTTTAAACCTGGCATCATTTCCCTCTGTAAAGCACATTTGGAAGACAGTACAGATACCT 2025
3079 TTTTAAACCTGGCATCATTTCCCTCTGTAAAGCACATTTGGAAGACAGTACAGATACCT 3138
2026 TTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCAGGCTGGGCTCTCTTCT 2085
3139 TTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCAGGCTGGGCTCTCTTCT 3198
2086 GCATGATTCATCCAA 2101
3199 GCATGATTCATCCAA 3214

RESULT 4
IS-09-845-416-10
Sequence 10, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 3531
TYPE: DNA
ORGANISM: Homo sapiens
IS-09-845-416-10

Query Match 83.3%; Score 1750.8; DB 13; Length 3531;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;
1 GAGCTATGCTTACACACAGGCTGCTTATGTGACCAACCTCTGACCCCTACACGAGCCCAT 60
900 GAGCTATGCTTACACACAGGCTGCTTATGTGACCAACCTCTGACCCCTACACGAGCCCAT 959
61 TCTTTCACAGCATTTGGAAGCTCTTGAAGACAGTCAATTTGGCAGTTCATTTGATGGAGAG 120
960 TCTTTCACAGCATTTGGAAGCTCTTGAAGACAGTCAATTTGGCAGTTCATTTGATGGAGAG 1019
121 TGAAGTAAACCTGGAGCCGTTATCAAAAGCTTTTGAAGAGTATTATCTGCTGCTCTTTC 180

Db 1020 TGAAGTAAACCTGGAGCCGTTATCAAAACAGCTTTAGAAGAAGTATTATCTGCTGCTCTTTC 1079
Qy 181 TGCTGAGGACACATTTGCAAGCAACAGGAGAGATTTCTTAATGATGTGGAAGTGTGGAAGA 240
Db 1080 TGCTGAGGACACATTTGCAAGCAACAGGAGAGATTTCTTAATGATGTGGAAGTGTGGAAGA 1139
Qy 241 CCAGTTTCATACCTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGCGCGGTGG 300
Db 1140 CCAGTTTCATACCTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGCGCGGTGG 1199
Qy 301 TAATATTCTACAATTTGGAGTAAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 1200 TAATATTCTACAATTTGGAGTAAAGCTGATTTGGAACAGGAAATTTATCAGAAGATGAAGA 1259
Qy 361 AACTGAAGTACAAGACAGATGAATCTCTTAATCAAGATGGGAATGCTCTCAGGCTAGC 420
Db 1260 AACTGAAGTACAAGACAGATGAATCTCTTAATCAAGATGGGAATGCTCTCAGGCTAGC 1319
Qy 421 TAGCATGGAAGAAACAAAGCAATTTTACATAGAGTT-----TTAATGGATCTCCAGATCA 474
Db 1320 TAGCATGGAAGAAACAAAGCAATTTTACATAGAGTTCTATAGAACTCATAGATTACTGCAACAGTCCCCCT 1379
Qy 475 GAAACTGAAAGAGTTTGAATGACTGCTTAAACAAAACAGGAAGAAACAAAGGAAATGGA 534
Db 1380 GGACCTGGAAGAGTTTCTTTCCTGCTGCTTACAGAGCTGAAACCACTGCCAATGTCTCTACA 1439
Qy 535 GGAAGAGCTCTTGGAGCTGATCTTGAAGACCTAAACCGCAAGTACAACAACATAGT 594
Db 1440 GGATGCTACCCGTAAAGAAAGGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAA 1499
Qy 595 GCCTTCAAGAGATCTAGAAACAGAAACAAGTCAAGGTCAATTTCTCCTCCTCAGTCTGCT 654
Db 1500 ACATGGCAGACCTCCAGGTGAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGA 1559
Qy 655 GGTAG-----TTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTTGGAGA 702
Db 1560 TGAAGAACAGCCAAAAATCTTGATGATCTCTGGAAGTTCCTGATGATGCTGCTGTTACA 1619
Qy 703 ACAACTTAAG-----TATTGGAGATCGATGGGCAACATCTGTAGTGGACAGAGACCG 759
Db 1620 AGACGCTTTGGATAACATGAATTCAGTGGAGTGAATTCGGAAGAAAGTCTCTCAACAT 1679
Qy 760 CTGGGTTCTTTTACAAGACAGTCTTGACCAAGTGGAGCGTCTGCACCTTTCTCTGAGGA 819
Db 1680 TAGGTCCTTCTTTGGAAGCCAGTCTTGACCAAGTGGAGCGTCTGCACCTTTCTCTGAGGA 1739
Qy 820 ACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGGACCTTATTGAGG 879
Db 1740 ACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAGGACCTTATTGAGG 1799
Qy 880 CGACTTTCCAGCAGTTCAGAAGCAGAACCATGTACATAGGGCTCTCAAGAGGGAATTTGAA 939
Db 1800 CGACTTTCCAGCAGTTCAGAAGCAGAACCATGTACATAGGGCTCTCAAGAGGGAATTTGAA 1859
Qy 940 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTGAGAAATTTCTGACAGAGCA 999
Db 1860 AACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTGAGAAATTTCTGACAGAGCA 1919
Qy 1000 GCCTTTGGAGGAGTACAGAAACTCTTACCAGAGCCCGCAGAGAGTCCCTCTGAGGAGAG 1059
Db 1920 GCCTTTGGAGGAGTACAGAAACTCTTACCAGAGCCCGCAGAGAGTCCCTCTGAGGAGAG 1979
Qy 1060 AGCCCAAGATGTCACTCGGCTTTCTAGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGA 1119
Db 1980 AGCCCAAGATGTCACTCGGCTTTCTAGAAAGCAGGCTGAGGAGGTCAATATCTGAGTGGGA 2039
Qy 1120 ARAATTTGAACCTGCACTCGGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT 1179
Db 2040 ARAATTTGAACCTGCACTCGGCTGACTGGCAGAGAAATAGATGAGACCTTGAAGACT 2099
Qy 1180 CCAGGAACCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTGCGCAAGCTGAGTGTAT 1239
Db 2100 CCAGGAACCTTCAAGAGGCCACGATGAGTGGACCTCAAGCTGCGCAAGCTGAGTGTAT 2159

QY 1240 CAAGGATCTTGGAGCCGCTGGGAGATCTCTCATTTGACATCTCTCCAGATCACTCTCGA 1299
Db 2160 CAAGGATCTTGGAGCCGCTGGGAGATCTCTCATTTGACATCTCTCCAGATCACTCTCGA 2219
QY 1300 GAAAGTCAAGGACATTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTCGACCGTCAA 1359
Db 2220 GAAAGTCAAGGACATTCGAGGAGAAATTCGCGCTCTGAAAGAGAACGTCGACCGTCAA 2279
QY 1360 TGACCTTGCTCCGAGCTTACACATTTGGGANTTCAGCTCTCACGGTATAACCTCAGCAC 1419
Db 2280 TGACCTTGCTCCGAGCTTACACATTTGGGANTTCAGCTCTCACGGTATAACCTCAGCAC 2339
QY 1420 TCTGGAAGACCTTGAACACAGATGGAAGCTCTGAGGTCGCGCTCGAGGACCGAGTCAG 1479
Db 2340 TCTGGAAGACCTTGAACACAGATGGAAGCTCTGAGGTCGCGCTCGAGGACCGAGTCAG 2399
QY 1480 GAGCTGTGANTGAAGCCACAGGACATTTGGTTCAGCATCTCAGACATTTCTTTCCAGTC 1539
Db 2400 GAGCTGTGANTGAAGCCACAGGACATTTGGTTCAGCATCTCAGACATTTCTTTCCAGTC 2459
QY 1540 TGTCCAGGTCCTCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 1599
Db 2460 TGTCCAGGTCCTCTGGGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA 2519
QY 1600 CGAGACTCAAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCAAGTCTTTAGC 1659
Db 2520 CGAGACTCAAACTTGTGGGACCATCCCAAAATGACAGAGCTTACCAAGTCTTTAGC 2579
QY 1660 TGACCTGTATATGTACATCTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCA 1719
Db 2580 TGACCTGTATATGTACATCTCAGCTTATAGGACTGCCATGAACTCCGAGACTGCA 2639
QY 1720 GAAGGCCCTTTCTTGGATCTCTTGGGCTGTCTGAGCTGTCTGATGCTCTTGGACCGCA 1779
Db 2640 GAAGGCCCTTTCTTGGATCTCTTGGGCTGTCTGAGCTGTCTGATGCTCTTGGACCGCA 2699
QY 1780 CAACCTCAAGCAAAATGACAGCCCATGATATCTGAGAGTATTAATTTGATGACAC 1839
Db 2700 CAACCTCAAGCAAAATGACAGCCCATGATATCTGAGAGTATTAATTTGATGACAC 2759
QY 1840 TATTTATGACCCCTGGAGCAAGACCAAAATTTGATGATACCGGACGATCCGCTGT 1899
Db 2760 TATTTATGACCCCTGGAGCAAGACCAAAATTTGATGATACCGGACGATCCGCTGT 1959
QY 1900 TATGTCGTGAATGCTGCTGAATTTTATGATACCGGACGATCCGCTGT 2079
Db 2820 TATGTCGTGAATGCTGCTGAATTTTATGATACCGGACGATCCGCTGT 2099
QY 1960 CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAGTACAG 2019
Db 2880 CTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACATTTGGAAGACAGTACAG 2039
QY 2020 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACGAGATTTTGTGACCAAGCCAGCTGGGCT 2079
Db 2940 ATACCTTTTCAAGCAAGTGGCAAGTTTCAACGAGATTTTGTGACCAAGCCAGCTGGGCT 2099
QY 2080 CCTTCTGATGATCTTATCCAA 2101
Db 3000 CCTTCTGATGATCTTATCCAA 3021

RESULT 5

US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE THEREOF
; FILE OF INVENTION: 2001-04-30
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/200,777

; PRIOR FILING DATE: 2000-04-28

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 30

; LENGTH: 4498

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-845-416-30

Query Match 83.3%; Score 1750.8; DB 13; Length 4498;

Best Local Similarity 90.2%; Pred. No. 0;

Matches 1914; Conservative 0; Mismatches 187; Indels 21; Gaps 3;

QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTACACACCTCTGACCCCTCAGCGAGGCCATT 60
Db 1657 GAGCTATGCTTACACACAGGCTGCTTATGTACACACCTCTGACCCCTCAGCGAGGCCATT 1716
QY 61 TCTTTCACAGCATTTGGAAGCTCTTGAAGACAAGTCATTTGSCAGTTTCTTATGATGAGAG 120
Db 1717 TCTTTCACAGCATTTGGAAGCTCTTGAAGACAAGTCATTTGSCAGTTTCTTATGATGAGAG 1776
QY 121 TGAAGTAAACCTTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTTATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTTGGACCGTTATCAAAACAGCTTTAGAAGAAGTATTTATCGTGGCTTCTTTC 1836
QY 181 TGTGTAGGACACATTCGAAGCAAGGAGAGATTTCTAATGATGGAAGTGGTGGTGAAGA 240
Db 1837 TGTGTAGGACACATTCGAAGCAAGGAGAGATTTCTAATGATGGAAGTGGTGGTGAAGA 1896
QY 241 CCAGTTTCATCTCATGAGGGGTATCATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1897 CCAGTTTCATCTCATGAGGGGTATCATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1956
QY 301 TAAATTTCTCAAAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAATGAAGA 360
Db 1957 TAAATTTCTCAAAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAAATGAAGA 2016
QY 361 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCTCAGGGTAGC 2076
QY 421 TAGCATGGAACCAAAAGCAATTTACATAGATTTTAAATGGATCTCTCAGAAATCA 474
Db 2077 TAGCATGGAACCAAAAGCAATTTACATAGATTTTAAATGGATCTCTCAGAAATCA 2136
QY 475 GAACTGAAAGAGTTGAATGATCTGGCTAACAAAAACAGAGAAAGAAACAGGAAATGGA 534
Db 2137 GGAACCTTGGAAAAAGTTTCTTGGCTTACAGAGCTGAAACCAACTGCTCAATGTCTTACA 2196
QY 535 GGAAGAGCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACACACATAAGGT 594
Db 2197 GGAATCTACCGGTAGGAAGAGCTCTTGAAGACCTTAAAGAGGTAAAGAGCTGATGAA 2256
QY 595 GCTTCAAGAGATCTTAGAACAGAAAGCAAGTCAGGGTCAATTTCTCTCACTCAATGGTGT 654
Db 2257 ACAATGGCAAGACCTTCAAGGTGAAATTTGAAGTCAACAGATGTTTATCACAACCTGGA 2316
QY 655 GGTAG-----TTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGGAGA 702
Db 2317 TGAAGAACAGCCAAAAAATCTGAGATCTCTGGAAGGTTCCGATGATGCTGCTTGTACA 2376
QY 703 ACAACTTAAAG-----TATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCG 759
Db 2377 AAGACGTTTGGATAACATGAATCTTCAAGTGGAGTGAATTCGMAAAAGTCTCTCAACAT 2436
QY 760 CTGGGTTCTTTTACAGACAGTCTGACCAAGTGGAGCGTCTGACCTTCTCTCAGGA 819
Db 2437 TAGGTCCTTCTTTGGAAGCCAGTTCTGACCAAGTGGAGCGTCTGACCTTCTCTCAGGA 2496
QY 820 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTTATGGAG 879
Db 2497 ACTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCTTATGGAG 2556

QY 880 CGACTTTCCAGCAGTTTCCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 939
DB 2557 CGACTTTCCAGCAGTTTCCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTGAA 2616
QY 940 AACTAAGAGACCTGTAATCATGAGTACTCTTGGAGCTGTACGATATTTCTGACAGACA 999
DB 2617 AACTAAGAGACCTGTAATCATGAGTACTCTTGGAGCTGTACGATATTTCTGACAGACA 2676
QY 1000 GCCTTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCCTCAGAGAG 1059
DB 2677 GCCTTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAGCTGCTCCTCAGAGAG 2736
QY 1060 AGCCAGAGATGTCTAGGCTTTTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA 1119
DB 2737 AGCCAGAGATGTCTAGGCTTTTACGAAAGCAGGCTGAGAGGTCAATACTGAGTGGGA 2796
QY 1120 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTCAAGAGACT 1179
DB 2797 AAAATTGAACCTGCACTCCGCTGACTGGCAGAGAAAATAGATGAGACCTTCAAGAGACT 2856
QY 1180 CCAGGAATTTCAAGAGCCACGAGTGGACCTCAAGCTGGCCAAAGCTGAGGTGAT 1239
DB 2857 CCAGGAATTTCAAGAGCCACGAGTGGACCTCAAGCTGGCCAAAGCTGAGGTGAT 2916
QY 1240 CAAGGGATCCTGGCAGCCGCTGGGCGATCTCTCATGACTCTCTCAAGATCACCTCGA 1299
DB 2917 CAAGGGATCCTGGCAGCCGCTGGGCGATCTCTCATGACTCTCTCAAGATCACCTCGA 2976
QY 1300 GAAAGTCAAGGCACTTCAGAGAGAGAAATGGCGCTCTGAAGAGAGAGCTGAGCCACCTCAA 1359
DB 2977 GAAAGTCAAGGCACTTCAGAGAGAGAAATGGCGCTCTGAAGAGAGAGCTGAGCCACCTCAA 3036
QY 1360 TGACCTTGCTGGCCAGCTTACACTTTGGGCACTTGGGCACTTACCGTATTAACCTCAGCAC 1419
DB 3037 TGACCTTGCTGGCCAGCTTACACTTTGGGCACTTGGGCACTTACCGTATTAACCTCAGCAC 3096
QY 1420 TCTGGAAGACTGAACACAGATGGAAGCTTCTGACAGTGGCGCTGAGGACCGAGTCA 1479
DB 3097 TCTGGAAGACTGAACACAGATGGAAGCTTCTGACAGTGGCGCTGAGGACCGAGTCA 3156
QY 1480 GCAGCTGCATGAAGCCCAAGGAGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTC 1539
DB 3157 GCAGCTGCATGAAGCCCAAGGAGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTC 3216
QY 1540 TGTCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 1599
DB 3217 TGTCCAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACCA 3276
QY 1600 CGAGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 1659
DB 3277 CGAGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGC 3336
QY 1660 TGACCTGAATTAATGTAGATTTCTCAGCTTATAGACTGCCATGAACTCGAAGACTGCA 1719
DB 3337 TGACCTGAATTAATGTAGATTTCTCAGCTTATAGACTGCCATGAACTCGAAGACTGCA 3396
QY 1720 GAAGGCCCTTTGCTTGGATCTCTTGAGCTGTGAGCTGATGTGATGCTTGACCCAGCA 1779
DB 3397 GAAGGCCCTTTGCTTGGATCTCTTGAGCTGTGAGCTGATGTGATGCTTGACCCAGCA 3456
QY 1780 CAACCTCAAGAAAATGACAGCCCATGGATATCTTGCAGATTAATTAATGTTTGACCA 1839
DB 3457 CAACCTCAAGAAAATGACAGCCCATGGATATCTTGCAGATTAATTAATGTTTGACCA 3516
QY 1840 TATTTATGACCGCTGGAGCAGAGCAGACAAATTTGGTCAAGCTCCCTCTCGGTGGA 1899
DB 3517 TATTTATGACCGCTGGAGCAGAGCAGACAAATTTGGTCAAGCTCCCTCTCGGTGGA 3576
QY 1900 TATGTGTCTGAACCTGGCTGTGATGTTTATGATACGGGACGAAACAGGAGGATCCGTCT 1959
DB 3577 TATGTGTCTGAACCTGGCTGTGATGTTTATGATACGGGACGAAACAGGAGGATCCGTCT 3636

QY 1960 CCTGTCTTTTAAACGTGCATCATTTCCCTGTGTAAAGCACAATTTGGAAGACAAGTACAG 2019
DB 3637 CCTGTCTTTTAAACGTGCATCATTTCCCTGTGTGTAAAGCACAATTTGGAAGACAAGTACAG 3696
QY 2020 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCCAGGCTGGGCT 2079
DB 3697 ATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAGCCAGGCTGGGCT 3756
QY 2080 CCTTCTGCATGATTTCTATCCAA 2101
DB 3757 CCTTCTGCATGATTTCTATCCAA 3778
RESULT 6
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3858
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-9
Query Match 83.0%; Score 1743; DB 13; Length 3858;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 GAGCTATGCTTACACACAGGCTGCTTATGTGTCACCACTCTGACCCCTACAGGAGCCCAT 60
DB 900 GAGCTATGCTTACACACAGGCTGCTTATGTGTCACCACTCTGACCCCTACAGGAGCCCAT 959
QY 61 TCCCTTCAGCAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 120
DB 960 TCCCTTCAGCAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTTCATTGATGGAGAG 1019
QY 121 TGAAGTAACTGACCGTTTATCAACAGCTTTTGAAGAAGTATTTATCGTGGCTTCTTC 180
DB 1020 TGAAGTAACTGACCGTTTATCAACAGCTTTTGAAGAAGTATTTATCGTGGCTTCTTC 1079
QY 181 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
DB 1080 TGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1139
QY 241 CCAAGTTTCACTACTCAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 300
DB 1140 CCAAGTTTCACTACTCAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTGG 1199
QY 301 TAATATTCTAATTTGGAAGTAACTGATTTGGAACAGGAAATTTATCAGAAGTGAAGA 360
DB 1200 TAATATTCTAATTTGGAAGTAACTGATTTGGAACAGGAAATTTATCAGAAGTGAAGA 1259
QY 361 AACTGAAGTCAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 420
DB 1260 AACTGAAGTCAAGAGCAGATGAATCTCTAAATTCAGATGGGAATGCCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTATCGATCTCCAGATCAGAACT 480
DB 1320 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTATCGATCTCCAGATCAGAACT 1379
QY 481 GAAAGAGTTGAATGACTGGCTTACAAAAACAGAAAGAAACAAGAAAAATGAGGAGAA 540
DB 1380 GAAAGAGTTGAATGACTGGCTTACAAAAACAGAAAGAAACAAGAAAAATGAGGAGAA 1439

QY 541 GCCTCTTGACCTGATCTTGAAGACCTAAACGCGCAAGTACAAACAATAGGTGCTTCA 600
DB 1440 GCCTCTTGACCTGATCTTGAAGACCTAAACGCGCAAGTACAAACAATAGGTGCTTCA 1499
QY 601 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTACTCAATGCGGTGAGT 660
DB 1500 AGAAGATCTAGAACAAAGAACAAAGTCAAGGTCAATCTCTACTCAATGCGGTGAGT 1559
QY 661 TGATCAATCTAGTGAGATCAGCAACTGCTGCTTGGAGAAACAATTAAGGTATTGGG 720
DB 1560 TGATCAATCTAGTGAGATCAGCAACTGCTGCTTGGAGAAACAATTAAGGTATTGGG 1619
QY 721 AGATCGATGGGCAACATCTGTAGTAGACAGACCGCTGGGTCTTTTACAGA --- 777
DB 1620 AGATCGATGGGCAACATCTGTAGTAGACAGACCGCTGGGTCTTTTACAGA --- 1679
QY 778 --- 777
DB 1680 TCATAGATTACTGCAACAGTTCCTCCCTGGAACCTGGAATAAAGTTCTTGCTGCTTACAGA 1739
QY 778 --- 777
DB 1740 ACCTGAACAACTGCCAATGTCTACAGATGCTACCCGTAAGGAAGGCTCCTTAGAAGA 1799
QY 778 --- 777
DB 1800 CTCCAAGGGTAAGAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAATTTGAAGC 1859
QY 778 --- 777
DB 1860 TCACACAGATGTTTATCAACACTGGATGAACACAGCCAAAATCTGAGATCCCTGGA 1919
QY 778 --- 777
DB 1920 AGTTCCGATGATGAGTCTGTTACAAAGAGCTTTGGATGAATGAATCAAGTCAAGTGGAG 1979
QY 778 --- 777
DB 1980 TGAATCTCGAAAAAGTCTCTCAACATPAGGTCCCATTTGGAAGCCAGTTCTGACCAGTG 2039
QY 793 GAAGCGCTGCAACCTTTCTCTCAGCAAACTTCTGGTGGCTCAGCTGAAAGATGATGA 852
DB 2040 GAAGCGCTGCAACCTTTCTCTCAGCAAACTTCTGGTGGCTCAGCTGAAAGATGATGA 2099
QY 853 ATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTCAGAGCAGACGATGT 912
DB 2100 ATTAAGCCGGCAGGACCTATTGGAGGCGACTTTCCAGCAGTTCAGAGCAGACGATGT 2159
QY 913 ACATAGGGCTTCAAGAGGGAATTCGAAACTGAAAGAACTGTAATCATGATGATCTTTGA 972
DB 2160 ACATAGGGCTTCAAGAGGGAATTCGAAACTGAAAGAACTGTAATCATGATGATCTTTGA 2219
QY 973 GACTGTACGAATATTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACAGGA 1032
DB 2220 GACTGTACGAATATTCTGACAGAGCAGCTTTGGAAGGACTAGAGAACTCTACAGGA 2279
QY 1033 GCCCAGAGAGCTGCTCTCCTGAGGAGGAGCCAGAACTGCTCTGCTGCTTCTACGAAGCA 1092
DB 2280 GCCCAGAGAGCTGCTCTCCTGAGGAGGAGCCAGAACTGCTCTGCTGCTTCTACGAAGCA 2339
QY 1093 GGCTGAGGAGGTCATPACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 1152
DB 2340 GGCTGAGGAGGTCATPACTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGGCAGAG 2399
QY 1153 AAAATAGATGAGACCTTTGAAAGACTCCAGAACTTCAAGGCGCAGGATGAGCTGGA 1212
DB 2400 AAAATAGATGAGACCTTTGAAAGACTCCAGAACTTCAAGGCGCAGGATGAGCTGGA 2459
QY 1213 CTTCAAGCTGCGCAGCTGAGTGAATCAAGGATCTTGGCAGCCGCTGGGCGATCTCTCT 1272
DB 2460 CTTCAAGCTGCGCAGCTGAGTGAATCAAGGATCTTGGCAGCCGCTGGGCGATCTCTCT 2519

QY 1273 CATTGACTCTCTCCAGATCACTTCGAGAAAGTCAAGCACTTCGAGGAGAAATTCGCC 1332
DB 2520 CATTGACTCTCTCCAGATCACTTCGAGAAAGTCAAGCACTTCGAGGAGAAATTCGCC 2579
QY 1333 TGTAAAAGAAAGTGTAGGCAAGTCAATGACCTTCTGCGCAGCTTACCACTTTGGGAT 1392
DB 2580 TGTAAAAGAAAGTGTAGGCAAGTCAATGACCTTCTGCGCAGCTTACCACTTTGGGAT 2639
QY 1393 TCAGCTCTCACCCTATTAACCTCAGCACTCTGGAAGACCTGAAACCAAGATGGAAGCTTCT 1452
DB 2640 TCAGCTCTCACCCTATTAACCTCAGCACTCTGGAAGACCTGAAACCAAGATGGAAGCTTCT 2699
QY 1453 GAGGTGGCGCTGAGGAGCCGAGTCAAGCAGTGTGATGAAGCCCAAGGAGCTTTGGTCC 1512
DB 2700 GAGGTGGCGCTGAGGAGCCGAGTCAAGCAGTGTGATGAAGCCCAAGGAGCTTTGGTCC 2759
QY 1513 AGCATCTCAGCACTTTCTTTCCACGCTGCTCCAGGCTCCCTGGGAGAGAGCCATCTCGCC 1572
DB 2760 AGCATCTCAGCACTTTCTTTCCACGCTGCTCCAGGCTCCCTGGGAGAGAGCCATCTCGCC 2819
QY 1573 AAACAAAGTCCCTACTATATCAACCAAGAGACTCAAAACAATCTGCTGGACCATCCCAA 1632
DB 2820 AAACAAAGTCCCTACTATATCAACCAAGAGACTCAAAACAATCTGCTGGACCATCCCAA 2879
QY 1633 AATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATATGTCAGATCTCAGCTTATAG 1692
DB 2880 AATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATATGTCAGATCTCAGCTTATAG 2939
QY 1693 GACTGCTCAAACTCCGAAAGTCTGCAAGAGCCCTTTGCTTGGATCTCTTGGAGCTGTC 1752
DB 2940 GACTGCTCAAACTCCGAAAGTCTGCAAGAGCCCTTTGCTTGGATCTCTTGGAGCTGTC 2999
QY 1753 AGCTGATGTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGGATAT 1812
DB 3000 AGCTGATGTGATGCTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGGATAT 3059
QY 1813 CTTGCAAGATTAATTAATGTTGACCACTATTATATGACCGCTTGGAGCAAGAGCAACAA 1872
DB 3060 CTTGCAAGATTAATTAATGTTGACCACTATTATATGACCGCTTGGAGCAAGAGCAACAA 3119
QY 1873 TTTGGTCAACGCTCCCTCTCTGCTGGATATGTCGAACTGGCTGCTGTAATGTTTATGA 1932
DB 3120 TTTGGTCAACGCTCCCTCTCTGCTGGATATGTCGAACTGGCTGCTGTAATGTTTATGA 3179
QY 1933 TAGGGACCAACAGAGGAGATCCGCTGCTGCTTTTAAACTGGCATCATTTCCCTGTG 1992
DB 3180 TAGGGACCAACAGAGGAGATCCGCTGCTGCTTTTAAACTGGCATCATTTCCCTGTG 3239
QY 1993 TAAAGCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTCAACAGG 2052
DB 3240 TAAAGCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTCAACAGG 3299
QY 2053 ATTTTGTGACCGGAGGCTGGGCTCTTCTGATGATCTATCCAA 2101
DB 3300 ATTTTGTGACCGGAGGCTGGGCTCTTCTGATGATCTATCCAA 3348

RESULT 7

US-09-845-416-29
; Sequence 29, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29

; LENGTH: 4825

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-845-416-29

Query Match 83.0%; Score 1743; DB 13; Length 4825;

Best Local Similarity 85.8%; Pred. No. 0;

Matches 2101; Conservative 0; Mismatches 0; Indels 346; Gaps 1;

```
2Y 1 GAGCTATGCTACACAGAGCTCTCTATATGTCACCACTCTGACCCCTACACGAGGCCATT 60
Db 1657 GAGCTATGCTACACAGAGCTCTCTATATGTCACCACTCTGACCCCTACACGAGGCCATT 1716
2Y 61 TCCTTCACAGCAATTTGGAAGCTCCTCAAGACAGTCAATTTGCGAGTTCAATGATGAGAG 120
Db 1717 TCCTTCACAGCAATTTGGAAGCTCCTCAAGACAGTCAATTTGCGAGTTCAATGATGAGAG 1776
2Y 121 TGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAGTATTTATCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAGTATTTATCGTGGCTTCTTTC 1836
2Y 181 TGCTGAGGACACATTTCAAGACCAAGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 240
Db 1837 TGCTGAGGACACATTTCAAGACCAAGAGAGATTTCTTAATGATGTGGAAGTGGTGAAGA 1896
2Y 241 CGAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1897 CGAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGGTGG 1956
2Y 301 TAATATTTACAAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 360
Db 1957 TAATATTTACAAATTTGGGAAGTAAGCTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 2016
2Y 361 AACTGAAGTACAGAGCAGATGAATCTCTTAATTAAGATCGGAATGCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAGAGCAGATGAATCTCTTAATTAAGATCGGAATGCTCAGGGTAGC 2076
2Y 421 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAACT 480
Db 2077 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAAATCAGAACT 2136
2Y 481 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAAACAAGAAAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGAAACAAGAAAAATGGAGGAAGA 2196
2Y 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAGGCTCTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTACAAACATAGGCTCTTCA 2256
2Y 601 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCACTCACTGGTGGTGTAGT 660
Db 2257 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCACTCACTGGTGGTGTAGT 2316
2Y 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAAACAACCTTAAGGTATTGGG 2376
2Y 721 AGATCGATGGGCAACATCTGTAGATGGAACAGACCGCTGGTCTTTTACAGA --- 777
Db 2377 AGATCGATGGGCAACATCTGTAGATGGAACAGACCGCTGGTCTTTTACAGA --- 2436
2Y 778 ----- 777
Db 2437 TCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAAAGTTTCTTGCTGCTTACAGA 2496
2Y 778 ----- 777
Db 2497 AGCTGAACAACTGCCAATGTCTACAGATGCTACCCGTAAGAAAGCTCCTAGAAGA 2556
2Y 778 ----- 777
Db 2557 CTCCAAGGGNGTAAGAGCTGATGAACAATGGCAAGACCTCCAAGGTGAATTTGAAGC 2616
```

```
QY 778 ----- 777
Db 2617 TCACACAGATGTTTATCACAACTGGATGAAGCCAAAAAATCTGAGATCCCTGGA 2676
QY 778 ----- 777
Db 2677 AGGTTCCGATGATGACAGTCTCTGTTACAAAGACGTTTGGATAACATGAACCTTCAAGTGGAG 2736
QY 778 ----- 777
Db 2737 TGAATTCGGAAAAAGTCTCTCAACATTTAGTCCCATTTTGGAGCCAGTCTCTGACCCAGTG 2796
QY 793 GAAGCGTCTGCACTTTCTCTCTGCAAGGAATCTTCTGGTGTGGCTACAGCTGAAGATGATGA 852
Db 2797 GAAGCGTCTGCACTTTCTCTCTGCAAGGAATCTTCTGGTGTGGCTACAGCTGAAGATGATGA 2856
QY 853 ATTAAGCCGCGAGCAGCCTATTTGGAGGCGAGCTTTCCAGCAGATTTCAAGAGCAGAACGATGT 912
Db 2857 ATTAAGCCGCGAGCAGCCTATTTGGAGGCGAGCTTTCCAGCAGATTTCAAGAGCAGAACGATGT 2916
QY 913 ACATPAGGCGCTTCAAGAGGGAATTTGAAACTAAAGAACTGTAAATCATGAGTACTCTTGA 972
Db 2917 ACATPAGGCGCTTCAAGAGGGAATTTGAAACTAAAGAACTGTAAATCATGAGTACTCTTGA 2976
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACCAAGCA 1032
Db 2977 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGGAGGACTAGAGAACTCTACCAAGCA 3036
QY 1033 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 1092
Db 3037 GCCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 3096
QY 1093 GGCTGAGAGGTCAATACTAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGGAGAG 1152
Db 3097 GGCTGAGAGGTCAATACTAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGGAGAG 3156
QY 1153 AAAAAATAGATGAGACCTTCAAGAGCTTCAAGAGGCTTCAAGAGGCGCCAGGATGAGCTGGA 1212
Db 3157 AAAAAATAGATGAGACCTTCAAGAGCTTCAAGAGGCTTCAAGAGGCGCCAGGATGAGCTGGA 3216
QY 1213 CCTCAAGCTCGGCAAGCTGAGTGATCAAGGATCTGCGACGCCGCTGGGCGATCTCT 1272
Db 3217 CCTCAAGCTCGGCAAGCTGAGTGATCAAGGATCTGCGACGCCGCTGGGCGATCTCTCT 3276
QY 1273 CATTTGACTCTCTCCAGATCACCTCCAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGC 1332
Db 3277 CATTTGACTCTCTCCAGATCACCTCCAGAAAGTCAAGGCACTTCGAGGAGAAATTCGCGC 3336
QY 1333 TCTGAAAGAGAGCTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCACCTTTGGGCAAT 1392
Db 3337 TCTGAAAGAGAGCTGAGCCACGTCATGACCTTGTCTGCGCAGCTTACCACCTTTGGGCAAT 3396
QY 1393 TCAGCTCTCAACGTTAACTCTGAGCACTCTGGAAGACCTGAAACACAGATGGAAGTTCT 1452
Db 3397 TCAGCTCTCAACGTTAACTCTGAGCACTCTGGAAGACCTGAAACACAGATGGAAGTTCT 3456
QY 1453 GCAGTGGCGCTGAGGACCGAGTCAAGGAGCTGATGAAGCCCAAGGAGCTTTGGTTC 1512
Db 3457 GCAGTGGCGCTGAGGACCGAGTCAAGGAGCTGATGAAGCCCAAGGAGCTTTGGTTC 3516
QY 1513 AGCATCTCAGCACTTTCTTTCCACGCTGTCTCAGGCTGCTGGGAGAGGCCATCTCGCC 1572
Db 3517 AGCATCTCAGCACTTTCTTTCCACGCTGTCTCAGGCTGCTGGGAGAGGCCATCTCGCC 3576
QY 1573 AAACAAAGTCCCTACTATATCAACCAAGAGCTCAAAACCTTGTCTGGGACCATCCCAA 1632
Db 3577 AAACAAAGTCCCTACTATATCAACCAAGAGCTCAAAACCTTGTCTGGGACCATCCCAA 3636
QY 1633 AATGACAGAGCTTACAGGCTTTTAGCTGACCTGAATAATGTGTCAGATTTCTAGCTTATAG 1692
Db 3637 AATGACAGAGCTTACAGGCTTTTAGCTGACCTGAATAATGTGTCAGATTTCTAGCTTATAG 3696
QY 1693 GACTGCCATGAAACTCGAAGACTGCAAGAGGCGCTTTGCTTGGATCTCTTGGACCTGTC 1752
```


Db 3697 GACTGCCATGAACCTCCGAGACTCAGAGGCCCTTGTGATCTCTTGAGCCCTGTC 3756
QY 1753 AGCTGATGTGATGCTTGGACAGACCAACTCAAGCAAAATGACAGCCCATGGATAT 1812
Db 3757 AGCTGATGTGATGCTTGGACAGACCAACTCAAGCAAAATGACAGCCCATGGATAT 3816
QY 1813 CTGCGAGATTATTAATTTGTTGACCACTATTATGACCGCTGGAGCAAGAGACCAAA 1872
Db 3817 CTGCGAGATTATTAATTTGTTGACCACTATTATGACCGCTGGAGCAAGAGACCAAA 3876
QY 1873 TTGTCACAGCTCTCTCTGCTGCTGATGATGCTGATGCTGATGCTGATGCTGATG 1932
Db 3877 TTGTCACAGCTCTCTCTGCTGCTGATGATGCTGATGCTGATGCTGATGCTGATG 3936
QY 1933 TAGGGACCAACAGGAGGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1992
Db 3937 TAGGGACCAACAGGAGGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3996
QY 1993 TAAAGCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 3997 TAAAGCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4056
QY 2053 ATTTTGTGACGAGCCAGGCTGGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2101
Db 4057 ATTTTGTGACGAGCCAGGCTGGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4105

RESULT 8
US-09-845-416-35
; Sequence 35, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 4848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-35

Query Match 83.0%; Score 1743; DB 13; Length 4848;
Best Local Similarity 85.8%; Pred No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;
QY 1 GAGCTATGCTTACACAGGCTGCTTATGTACACCTCTGACCCCTACACGAGCCCAT 60
Db 1680 GAGCTATGCTTACACAGGCTGCTTATGTACACCTCTGACCCCTACACGAGCCCAT 1739
QY 61 TCCTTACAGCATTTGGAGCTCCTGAGACAGTCAATTTGSCAGTTTCATTTGATGGAGAG 120
Db 1740 TCCTTACAGCATTTGGAGCTCCTGAGACAGTCAATTTGSCAGTTTCATTTGATGGAGAG 1799
QY 121 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAGATTTATCGTGGCTTTCTTC 180
Db 1800 TGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAGATTTATCGTGGCTTTCTTC 1859
QY 181 TGCTGAGGACATTTGACAGCAGAGAGATTTCTAATGATGTTGAAGTGGTGAAGA 240
Db 1860 TGCTGAGGACATTTGACAGCAGAGAGATTTCTAATGATGTTGAAGTGGTGAAGA 1919
QY 241 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1920 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1979

QY 301 TAATATTCTACAAATGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAATCAAGA 360
Db 1980 TAATATTCTACAAATGGGAAGTAAGCTGATTGGAACAGGAAATTTATCAGAAATCAAGA 2039
QY 361 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATCAAGATGGGAATGCCTCAGGTAGC 420
Db 2040 AACTGAAGTACAAGAGCAGATGAATCTCTCTAAATCAAGATGGGAATGCCTCAGGTAGC 2099
QY 421 TAGCATGGAAAAACAAGCAATTTTACATAGAGTCTTAAATGGATCTCCAGAATCAGAACT 480
Db 2100 TAGCATGGAAAAACAAGCAATTTTACATAGAGTCTTAAATGGATCTCCAGAATCAGAACT 2159
QY 481 GAAAGAGTTTCAATGACTGGCTAACAAAAACAAGAAAGAAACAAGAAATGGAGGAAGA 540
Db 2160 GAAAGAGTTTCAATGACTGGCTAACAAAAACAAGAAAGAAACAAGAAATGGAGGAAGA 2219
QY 541 GCCTCTTGACCTGATCTTTGAGAGCTTAAAGCCCAAGTACCAACATCAAGCTGCTTCA 600
Db 2220 GCCTCTTGACCTGATCTTTGAGAGCTTAAAGCCCAAGTACCAACATCAAGCTGCTTCA 2279
QY 601 AGAAGATCTAGAACCAAGAAACAAGTCAAGGTCAATTTCTCTCACTCACTGCTGCTGCT 660
Db 2280 AGAAGATCTAGAACCAAGAAACAAGTCAAGGTCAATTTCTCTCACTCACTGCTGCTGCT 2339
QY 661 TGAATGATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAGCAACTTAAAGTATGGG 720
Db 2340 TGAATGATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAGCAACTTAAAGTATGGG 2399
QY 721 AGATCGATGGCAAAACATCTGTAGATGGAAGACCGCTGGTCTTTTACAAGA --- 777
Db 2400 AGATCGATGGCAAAACATCTGTAGATGGAAGACCGCTGGTCTTTTACAAGACAC 2459
QY 778 ----- 777
Db 2460 TCATAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAGATTTCTTGCTGGCTTACAGA 2519
QY 778 ----- 777
Db 2520 AGCTGAAACAACCTGCCAATGTCTTACAGGATGCTACCCGTAGGAAAGGCTCCTAGAAGA 2579
QY 778 ----- 777
Db 2580 CTCGAAGGAGTAAAGAGCTGATGAAACAATGGCAAGACCTTCCAAGGTGAATTTGAAGC 2639
QY 778 ----- 777
Db 2640 TCACACAGATGTTTATCAACACTGGATGAAACAAGCCAAAATAATCTTGAGATCCCTGGA 2699
QY 778 ----- 777
Db 2700 AGTTTCCGATGATGACAGTCTCTGTACAAAGACGTTTGGATACATGAATCTCAAGTGGAG 2759
QY 778 ----- 777
Db 2760 TGAATCTCGGAAAAAGTCTCTCAACATTAGTGTCCCATTTGGAGCCAGTCTTGACAGGTG 2819
QY 793 GAAGCGCTCTGCACCTTTCTCTGACAGAACTTCTGCTGTGGCTTACAGCTGAAAGATGATGA 852
Db 2820 GAAGCGCTCTGCACCTTTCTCTGACAGAACTTCTGCTGTGGCTTACAGCTGAAAGATGATGA 2879
QY 853 ATTAAGCCGCGGACGACCTTATTTGGAGCGGACCTTCCAGCAGTTTCAAGAGCAGAACGATGT 912
Db 2880 ATTAAGCCGCGGACGACCTTATTTGGAGCGGACCTTCCAGCAGTTTCAAGAGCAGAACGATGT 2939
QY 913 ACATAGGGCTTCAAGAGGGAATTTGAAATAAGAACTGTAATCATGAGTACTCTTGA 972
Db 2940 ACATAGGGCTTCAAGAGGGAATTTGAAATAAGAACTGTAATCATGAGTACTCTTGA 2999
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGAGCTAGAGAACTCTTACAGGA 1032
Db 3000 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGAGCTAGAGAACTCTTACAGGA 3059
QY 1033 GCCCAGAGAGCTGCCTCTCTGAGGAGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAAGCA 1092

Db 3060 GCCCAGAGAGCTCCCTCCTGAGGAGAGCCGAGAATGTCACTCGCTTCTACGAAGCA 3119
Qy 1093 GGTGTGAGAGGTCAATACTGAGTGGGAAAATTAACCTGCACTCCGCTGACTGGCAGAG 1152
Db 3120 GGTGTGAGAGGTCAATACTGAGTGGGAAAATTAACCTGCACTCCGCTGACTGGCAGAG 3179
Qy 1153 AAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGA 1212
Db 3180 AAAAATAGATGAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAGCTGA 3239
Qy 1213 CCTCAGCTCGCCAGCTGAGTGTATCAAGGATCCTGCGACCGCTGGGAGCTCCT 1272
Db 3240 CCTCAGCTCGCCAGCTGAGTGTATCAAGGATCCTGCGACCGCTGGGAGCTCCT 3299
Qy 1273 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGC 1332
Db 3300 CATTGACTCTCTCCAAAGATCACCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTCGCGC 3359
Qy 1333 TCTGAAAGAGAGCTGAGCCACGTCATGACCTTGTCTCGCAGCTTACCACTTTGGGAT 1392
Db 3360 TCTGAAAGAGAGCTGAGCCACGTCATGACCTTGTCTCGCAGCTTACCACTTTGGGAT 3419
Qy 1393 TCAGCTCTCACCGTATAAACCCTCAGCACTCTGGAAGACCTGAACACCAAGTGAAGCTTCT 1452
Db 3420 TCAGCTCTCACCGTATAAACCCTCAGCACTCTGGAAGACCTGAACACCAAGTGAAGCTTCT 3479
Qy 1453 GCAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGAAGCCCAAGGAGCTTTGGTCC 1512
Db 3480 GCAGGTGGCCGTCGAGGACCGAGTCAGGAGCTGCATGAAGCCCAAGGAGCTTTGGTCC 3539
Qy 1513 AGCATCTCAGCACTTTCTTCCACGCTCTGTCAGGCTCCCTGGGAGAGCCATCTCGCC 1572
Db 3540 AGCATCTCAGCACTTTCTTCCACGCTCTGTCAGGCTCCCTGGGAGAGCCATCTCGCC 3599
Qy 1573 AAAAAGTGCCTTACTATATCAACCAAGAGCTCAAAACCTTGTGGGACCATCCCAA 1632
Db 3600 AAAAAGTGCCTTACTATATCAACCAAGAGCTCAAAACCTTGTGGGACCATCCCAA 3659
Qy 1633 AATGACAGAGCTCTACAGCTTTAGTGACCTGAATATATGACATTTCTAGCTTTAG 1692
Db 3660 AATGACAGAGCTCTACAGCTTTAGTGACCTGAATATATGACATTTCTAGCTTTAG 3719
Qy 1693 GACTGCCATGAAGCTCCGAGAGCTGAGAGGCTTTGCTTGGATCTCTGAGCTGTC 1752
Db 3720 GACTGCCATGAAGCTCCGAGAGCTGAGAGGCTTTGCTTGGATCTCTGAGCTGTC 3779
Qy 1753 AGCTGATGTATGCTTGGACAGCAACCTCAAGGCAAAATGACCCAGCCATGGATAT 1812
Db 3780 AGCTGATGTATGCTTGGACAGCAACCTCAAGGCAAAATGACCCAGCCATGGATAT 3839
Qy 1813 CCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGACAA 1872
Db 3840 CCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGACAA 3899
Qy 1873 TTTGGTCAACGCTCCCTCTCTGCTGGATATGCTCTGAATGCTGCTGCTGAATGTTATGA 1932
Db 3900 TTTGGTCAACGCTCCCTCTCTGCTGGATATGCTCTGAATGCTGCTGCTGAATGTTATGA 3959
Qy 1933 TACGGGACGACAGGAGGATCCGCTCTGCTGCTTTTAAACTGGCATCTTCCCTGTG 1992
Db 3960 TACGGGACGACAGGAGGATCCGCTCTGCTGCTTTTAAACTGGCATCTTCCCTGTG 4019
Qy 1993 TAAAGCACTTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 2052
Db 4020 TAAAGCACTTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGG 4079
Qy 2053 ATTTTGTGACAGCGAGGCTGGGCTCTCTTCTGATGATTTCTATCCAA 2101
Db 4080 ATTTTGTGACAGCGAGGCTGGGCTCTCTTCTGATGATTTCTATCCAA 4128

US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match 83.0%; Score 1743; DB 13; Length 5060;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 348; Gaps 1;

Qy 1 GAGCTATGCGCTACACACAGAGCTGCTTATGTACACCTCTGACCCTACACGAGCCCAT 60
Db 1892 GAGCTATGCGCTACACACAGAGCTGCTTATGTACACCTCTGACCCTACACGAGCCCAT 1951
Qy 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCTATTTGGCAGTTTATTGATGGAGAG 120
Db 1952 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCTATTTGGCAGTTTATTGATGGAGAG 2011
Qy 121 TGAAGTAACTGACCGTTATCAACAGCTTTAGAGAGATTTATTCGTGGCTTCTTTC 180
Db 2012 TGAAGTAACTGACCGTTATCAACAGCTTTAGAGAGATTTATTCGTGGCTTCTTTC 2071
Qy 181 TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA 240
Db 2072 TGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTAATGATGTGGAAGTGTGAAGA 2131
Qy 241 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 300
Db 2132 CCAGTTTCTACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGTGG 2191
Qy 301 TAATATTTCTCAATTTGGAAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 360
Db 2192 TAATATTTCTCAATTTGGAAAGTAAAGCTGATTGGAACAGGAAATTTATCAGAAGATGAAGA 2251
Qy 361 AACTGAAGTCAAGAGCAGATGAATCTCTTAAATTCAGATGGGAATGCTCAGGGTAGC 420
Db 2252 AACTGAAGTCAAGAGCAGATGAATCTCTTAAATTCAGATGGGAATGCTCAGGGTAGC 2311
Qy 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 480
Db 2312 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 2371
Qy 481 GAAAGAGTTGAATGACCTGCTTAAACAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 540
Db 2372 GAAAGAGTTGAATGACCTGCTTAAACAAAACAGAAAGAAACAAAGGAAATGGAGGAAGA 2431
Qy 541 GCCTCTTGAGCCTGATCTTGAAGCCTTAAACCGCAAGTACAAACATAGGTGCTTCA 600
Db 2432 GCCTCTTGAGCCTGATCTTGAAGCCTTAAACCGCAAGTACAAACATAGGTGCTTCA 2491
Qy 601 AGAAGATCTAGAACAAACAAAGTCAGGGTCAATTTCTCTCACTCACATGCTGGTGGTAGT 660
Db 2492 AGAAGATCTAGAACAAACAAAGTCAGGGTCAATTTCTCTCACTCACATGCTGGTGGTAGT 2551
Qy 661 TGATGAATCTAGTGGAGATCAACGCAACTGCTGCTTTGGAAAGAAACAATTAAGGTATTGGG 720
Db 2552 TGATGAATCTAGTGGAGATCAACGCAACTGCTGCTTTGGAAAGAAACAATTAAGGTATTGGG 2611
Qy 721 AGATCGATGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAGA --- 777

Db 2612 AGATCGATGGGCAACAATCTGTAGATGACAGAGACCGCTGGGTTCTTTTACAAGACAC 2671
QY 778 ----- 777
Db 2672 TCATAGATTACTGCAACAGTTCCCGCTGGACCTGGAAAAAGTTTCTTGCCCTGGCTTACAGA 2731
QY 778 ----- 777
Db 2732 AGCTGAACAACCTGCCAATGCTCTACAGATGCTACCCGTGAAGAAAGGCTCCTAGAGA 2791
QY 778 ----- 777
Db 2792 CTCCAAGGAGTAAAGAGCTGATGAACAATGGCAAGACCTCAAGGTGAATTTGAAGC 2851
QY 778 ----- 777
Db 2852 TCACACAGATGTTTATCAACCTGGAAGAAACAGCCAAATAATCTGAGATCCCTGGA 2911
QY 778 ----- 777
Db 2912 AGTTCCGATGATGCAGTCTCTGTACAAAGAGCTTTGGATTAACATGAACCTTCAAGTGGAG 2971
QY 778 ----- 777
Db 2972 TGAACCTTCGGAAAAAGTCTCTCAACATTAGTGTCCCATTTGGAAGCCAGTTCTGACCAAGT 3031
QY 793 GAAGCGTCTGCACCTTTCTCTCAGGAACCTCTGGTGTGGCTACAGCTGAAAGATGATGA 852
Db 3032 GAAGCGTCTGCACCTTTCTCTCAGGAACCTCTGGTGTGGCTACAGCTGAAAGATGATGA 3091
QY 853 ATTAAAGCGGAGGACCTATTATGAGGAGGACCTTTCCAGAGTTTCAAGAGCAGACGATGT 912
Db 3092 ATTAAAGCGGAGGACCTATTATGAGGAGGACCTTTCCAGCAGTTTCAAGAGCAGACGATGT 3151
QY 913 ACATAGGCGCTTCAAGAGGGAATTTGAAGAACTAAAGAACTGTAATCATGACTCTTTGA 972
Db 3152 ACATAGGCGCTTCAAGAGGGAATTTGAAGAACTAAAGAACTGTAATCATGACTCTTTGA 3211
QY 973 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACCGGA 1032
Db 3212 GACTGTACGAATATTTCTGACAGAGCAGCCTTTTGAAGGACTAGAGAACTCTACCGGA 3271
QY 1033 GCCCAGAGAGCTGCTCTCTGAGGAGAGCCAGAGATGTCCTCGGCTTCTACGAAAGCA 1092
Db 3272 GCCCAGAGAGCTGCTCTCTGAGGAGAGCCAGAGATGTCCTCGGCTTCTACGAAAGCA 3331
QY 1093 GGCTGAGGAGGTCATATCTAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAG 1152
Db 3332 GGCTGAGGAGGTCATATCTAGTGGGAAAAATTTGAACCTGCACTCCGCTGACTGGCAGAG 3391
QY 1153 AAAAAATAGATGAGACCTTTGAAGACTCAGGAACCTTCAAGAGGCCAGGATGAGCTGGA 1212
Db 3392 AAAAAATAGATGAGACCTTTGAAGACTCAGGAACCTTCAAGAGGCCAGGATGAGCTGGA 3451
QY 1213 CTTCAAGCTGCGCAAGCTGAGTGTATCAAGGATCTCTGGCAGCCGCTGGGCGATCTCCT 1272
Db 3452 CTTCAAGCTGCGCAAGCTGAGTGTATCAAGGATCTCTGGCAGCCGCTGGGCGATCTCCT 3511
QY 1273 CATTTACTCTCTCAAGATCACTCTGAGAGAAAGTCAAGGCACTTCGAGAGAAATTTGCC 1332
Db 3512 CATTTACTCTCTCAAGATCACTCTGAGAGAAAGTCAAGGCACTTCGAGAGAAATTTGCC 3571
QY 1333 TCTGAAGAGAGAGCTGAGCCAGCTCAATGACCTTGTCCGACGCTTACCACTTTGGGAT 1392
Db 3572 TCTGAAGAGAGAGCTGAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCACTTTGGGAT 3631
QY 1393 TCAGCTCTCTCAACCGTATAAAGCTCAGCACTCTGGAAGACCTGGAACACAGATGAAGCTTCT 1452
Db 3632 TCAGCTCTCTCAACCGTATAAAGCTCAGCACTCTGGAAGACCTGGAACACAGATGAAGCTTCT 3691
QY 1453 GCAGGTGGCGCTGAGAGCCGAGTCAAGGAGCTGATGAAGCCACAGGACTTTGGTCC 1512

Db 3692 GCAGGTGGCGCTGAGAGCCGAGTCAAGCAGCTGATGAAGCCACAGGACTTTGGTCC 3751
QY 1513 AGCATCTCAGCACTTTCTTTTCCAGCTGTGTCCAGGTCCTCGGAGAGAGCCATCTCGCC 1572
Db 3752 AGCATCTCAGCACTTTCTTTTCCAGCTGTGTCCAGGTCCTCGGAGAGAGCCATCTCGCC 3811
QY 1573 AAACAAAGTGGCCTACTATATCAACACAGAGACTCAAAACAACTTCTGCGGACCATCCCAA 1632
Db 3812 AAACAAAGTGGCCTACTATATCAACACAGAGACTCAAAACAACTTCTGCGGACCATCCCAA 3871
QY 1633 AATGACAGAGCTCTACAGCTTTTACAGTCACTGACCTGAATATGTACAGTTCTCAGCTTATAG 1692
Db 3872 AATGACAGAGCTCTACAGCTTTTACAGTCACTGACCTGAATATGTACAGTTCTCAGCTTATAG 3931
QY 1693 GACTGCATGAATAACTCCGAGAGACTGCAGAGGCGCTTTGCTTGGATCTCTTGAAGCTGTC 1752
Db 3932 GACTGCATGAATAACTCCGAGAGACTGCAGAGGCGCTTTGCTTGGATCTCTTGAAGCTGTC 3991
QY 1753 AGTGCATGTGATGCTTGGACAGCAACCTCAAGCAAAATGACCCAGCCCATGATAT 1812
Db 3992 AGTGCATGTGATGCTTGGACAGCAACCTCAAGCAAAATGACCCAGCCCATGATAT 4051
QY 1813 CTTGCAAGATTTAATTTGTTGACCACTATTTATGACGCGCTGGAGCAAGACACAACAA 1872
Db 4052 CTTGCAAGATTTAATTTGTTGACCACTATTTATGACGCGCTGGAGCAAGACACAACAA 4111
QY 1873 TTTGGTCAACGTCCTCTCTGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGA 1932
Db 4112 TTTGGTCAACGTCCTCTCTGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGA 4171
QY 1933 TAGGAGCAAGACAGGAGGATCGGTGCTGCTCTTTTAAACTGGCATCATTTCCCTGTG 1992
Db 4172 TAGGAGCAAGACAGGAGGATCGGTGCTGCTCTTTTAAACTGGCATCATTTCCCTGTG 4231
QY 1993 TAAAGCACATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAG 2052
Db 4232 TAAAGCACATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAG 4291
QY 2053 ATTTGTGACAGCGAGCTGGGCGCTCTTCTGATGATTTCTATCCAA 2101
Db 4292 ATTTGTGACAGCGAGCTGGGCGCTCTTCTGATGATTTCTATCCAA 4340

RESULT 10

US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match 78.1%; Score 1641; DB 13; Length 4414;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1848; Conservative 0; Mismatches 190; Indels 63; Gaps 3;
QY 1 GAGCTATGCTCACACAGAGCTCTTATGTTCACCACTCTGACCTCTACGAGGCCATT 60
Db 1657 GAGCTATGCTCACACAGAGCTCTTATGTTCACCACTCTGACCTCTACGAGGCCATT 1716
QY 61 TCCTTACAGCATTTTGAAGCTCTCTGAAGACAGTCAATTTGGCAGTTTCATTGATGGAGAG 120

Db 1717 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAATTTGGCAGTTCAATGATGAGAG 1776
Qy 121 TGAAGTAAACCTGGACCGTTTATCAACAGCCTTTAGAGAGATTAATTCGTGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGGACCGTTTATCAACAGCCTTTAGAGAGATTAATTCGTGCTTCTTTC 1836
Qy 181 TGTGAGGACACATTTGCAAGCAACAAGAGAGATTTCTTAATGATGTGGAAGTGTGTAAGA 240
Db 1837 TGTGAGGACACATTTGCAAGCAACAAGAGAGATTTCTTAATGATGTGGAAGTGTGTAAGA 1896
Qy 241 CCAGTTTCACTACATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGGTTGG 300
Db 1897 CCAGTTTCACTACATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGCGGGTTGG 1956
Qy 301 TAATATTTCTACAAATGGGAAGTAAAGTGTGGAACAGGAAATTAATCAGAAAGATGAAGA 360
Db 1957 TAATATTTCTACAAATGGGAAGTAAAGTGTGGAACAGGAAATTAATCAGAAAGATGAAGA 2016
Qy 361 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCTCAGGGTAGC 420
Db 2017 AACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCTCAGGGTAGC 2076
Qy 421 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAACTCAGAAACT 480
Db 2077 TAGCATGGAAAAACAAGCAATTTACATAGATTTTAAATGGATCTCCAGAACTCAGAAACT 2136
Qy 481 GAAAGAGTTGAATGACTGGCTAACAACAAAGAGAAAGAAACAAGGAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTAACAACAAAGAGAAAGAAACAAGGAAATGGAGGAAGA 2196
Qy 541 GCCTCTTGGACCTGATCTTGAAGCACTTAAACCCCAAGTACACACATAAAGGTCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGCACTTAAACCCCAAGTACACACATAAAGGTCTTCA 2256
Qy 601 AGAAGATCTAGAACAAAGAAACAAGTCAAGGTCAATCTCTCACTCAGATGGTGGTAGT 660
Db 2257 AGAAGATCTAGAACAAAGAAACAAGTCAAGGTCAATCTCTCACTCAGATGGTGGTAGT 2316
Qy 661 TGATGAATCTAGTGGAGATCAGCAGACTGCTTTGGAGAGAACACTTAAGGTATTGGG 720
Db 2317 TGATGAATCTAGTGGAGATCAGCAGACTGCTTTGGAGAGAACACTTAAGGTATTGGG 2376
Qy 721 AGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGACAG 780
Db 2377 AGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGCTGGGTCTTTTACAAGACAT 2436
Qy 781 TTCTGACCACTGGAAGGCTGTGACACCTTTCTCTGACAGGAACCTTCTGGTGTGCTACAGCT 840
Db 2437 CCTTCTCAAAATGGCAACGCTTACTGAAGAACAGTGCCTTTTGTAGTGCATGGCTTTCAGA 2496
Qy 841 GAAAGATGATGAATTAAGCCGCGCAGGACCTATTTGGAGCGGACTTTCCAGCAGTTTCAGAA 900
Db 2497 AAGAGAGATGCACTGACACAGATTCACACACTGGCTTTAAGATCAAAATGAATGTT 2556
Qy 901 GCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACTAAGAACCTGTAAATCAT 960
Db 2557 ATCAAGTCTTCAAAACCTGGCCGCTTTTAAAGCGGATCTAGAAAAGAAAAAGCAATCCAT 2616
Qy 961 GAGTACTCTTGACACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAGAGGACTAGAGAA 1020
Db 2617 GGGCAAACTGTA-----TTCACTCAAAACAAGATCTTTCTTTCAACACTGAAG 2662
Qy 1021 ACTCTACAGAGCCCGCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCT 1080
Db 2663 AATAAGTCAGTGACCCAGAGAGCGAGCATGGCTGGATAACTTTGCCCGGTGTTGGG--- 2720
Qy 1081 TCTACGAAAGCAGGCTGAGAGGTCATTAATCTAGTGGGAAAAAATTTGAACCTGCACTCCGC 1140
Db 2721 -----ATAATTTAGTCCAAAAAATTTGAA----- 2743
Qy 1141 TGACTGSCAGAGAAAAATAGATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGGCCAC 1200

Db 2744 -----AAGAGTACAGCACAGACCCCTTTGAAGACTCCAGGAACCTTCAAGAGGCCAC 2793
Qy 1201 GGATGAGCTGAGACCTCAAGCTGCGCCAGCTGAGTGAATCAAGGGATCCTTGGCAGCCCGT 1260
Db 2794 GGATGAGCTGAGACCTCAAGCTGCGCCAGCTGAGTGAATCAAGGGATCCTTGGCAGCCCGT 2853
Qy 1261 GGGCGATCTCCTCAATTAATGATCTCTCTCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGG 1320
Db 2854 GGGCGATCTCCTCAATTAATGATCTCTCTCAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGG 2913
Qy 1321 AGAATTTGGCGCTCTGAAGAGACGTCGAGCCAGTCAATGACCTTCTGCTGCCAGCTTAC 1380
Db 2914 AGAATTTGGCGCTCTGAAGAGACGTCGAGCCAGTCAATGACCTTCTGCTGCCAGCTTAC 2973
Qy 1381 CACTTTGGCGATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTTGAAACACAG 1440
Db 2974 CACTTTGGCGATTCAGCTCTCACCGTATAACCTCAGCACCTCTGGAAGACCTTGAAACACAG 3033
Qy 1441 ATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGACGTCAGGACGTCATGAAGCCACAG 1500
Db 3034 ATGGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGACGTCATGAAGCCACAG 3093
Qy 1501 GGAATTTGGTCCAGCTCTCAGCACCTTTCTTCCAGCTCTGTCCAGGGTCCCTGGGAGAG 1560
Db 3094 GGAATTTGGTCCAGCTCTCAGCACCTTTCTTCCAGCTCTGTCCAGGGTCCCTGGGAGAG 3153
Qy 1561 AGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACACGAGACTCAAAACAATTTGCTG 1620
Db 3154 AGCCATCTCGCCAAACAAAGTGGCCCTACTATATCAACACGAGACTCAAAACAATTTGCTG 3213
Qy 1621 GGACCATCCCAAAATGACAGACTCTACAGCTCTTACAGCTCTTACGCTGCTGAATATGTCAGATT 1680
Db 3214 GGACCATCCCAAAATGACAGACTCTACAGCTCTTACAGCTCTTACGCTGCTGAATATGTCAGATT 3273
Qy 1681 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAAGAGCCCTTTGCTGGATCT 1740
Db 3274 CTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAAGAGCCCTTTGCTGGATCT 3333
Qy 1741 CTTGAGCCTGTACGCTGCATGTGATGCTTTGACCACTATTATGACCGCCTGGAGCA 1800
Db 3334 CTTGAGCCTGTACGCTGCATGTGATGCTTTGACCACTATTATGACCGCCTGGAGCA 3393
Qy 1801 GCCCATGCAATCTCTGAGATTAATTAATTTGACCACTATTATGACCGCCTGGAGCA 1860
Db 3394 GCCCATGCAATCTCTGAGATTAATTAATTTGACCACTATTATGACCGCCTGGAGCA 3453
Qy 1861 AGAGCAACAATTTGGTCAACGCTCCTCTCTCGGTGATATGTGCTGAATCTGGCTGCT 1920
Db 3454 AGAGCAACAATTTGGTCAACGCTCCTCTCTCGGTGATATGTGCTGAATCTGGCTGCT 3513
Qy 1921 GAATGTTTTATGATACGGGACGAAACAGGAGGATCCGCTGCTCTTTTAAACTGGCAT 1980
Db 3514 GAATGTTTTATGATACGGGACGAAACAGGAGGATCCGCTGCTCTTTTAAACTGGCAT 3573
Qy 1981 CATTTCCCTGTGTAAGCACTTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGC 2040
Db 3574 CATTTCCCTGTGTAAGCACTTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGC 3633
Qy 2041 AAGTTCAACAGGATTTTGTGACCGAGCAGCTGGGCTCCTTCTGCATGATTTATCCA 2100
Db 3634 AAGTTCAACAGGATTTTGTGACCGAGCAGCTGGGCTCCTTCTGCATGATTTATCCA 3693
Qy 2101 A 2101
Db 3694 A 3694

RESULT 11
US-09-845-416-14
; Sequence 14, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO

1 ; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

2 ; TITLE OF INVENTION: THEREOF

3 ; FILE REFERENCE: DE1142

4 ; CURRENT APPLICATION NUMBER: US/09/845,416

5 ; PRIOR FILING DATE: 2001-04-30

6 ; PRIOR APPLICATION NUMBER: 60/200,777

7 ; PRIOR FILING DATE: 2000-04-28

8 ; NUMBER OF SEQ ID NOS: 36

9 ; SOFTWARE: Patensin Ver. 2.1

10 ; SEQ ID NO 14

11 ; LENGTH: 3446

12 ; TYPE: DNA

13 ; ORGANISM: Homo sapiens

14 ; US-09-845-416-14

Query Match 77.5%; Score 1629; DB 13; Length 3446;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 1847; Conservative 0; Mismatches 190; Indels 64; Gaps 4;

```
QY 1 GAGCTATGCTACACACAGGCTGCTTATGTCCACCACTCTGACCCCTACACGAGCCCAATT 60
DB 900 GAGCTATGCTACACACAGGCTGCTTATGTCCACCACTCTGACCCCTACACGAGCCCAATT 959
QY 61 TCCCTTCACAGCATTTGGAGCTCTTGAAGACAGTCAATTTGGCAGTTCATTGATGGAGAG 120
DB 960 TCCCTTCACAGCATTTGGAGCTCTTGAAGACAGTCAATTTGGCAGTTCATTGATGGAGAG 1019
QY 121 TGAAGTAAACCTGGACCGCTTATCAACAGCTTTTGAAGAAGTATTATCGTGGCTTCTTTTC 180
DB 1020 TGAAGTAAACCTGGACCGCTTATCAACAGCTTTTGAAGAAGTATTATCGTGGCTTCTTTTC 1079
QY 181 TGCTGAGGACACATTTGGAAGTAACTGATGGAGAGAGATTTCTTAATGATGGAGTGGTGAAGA 240
DB 1080 TGCTGAGGACACATTTGGAAGTAACTGATGGAGAGAGATTTCTTAATGATGGAGTGGTGAAGA 1139
QY 241 CCAGTTTTCATCTCATGAGGGGTGACATGATGATTTGACGCCATCAGGGCCGGGTGG 300
DB 1140 CCAGTTTTCATCTCATGAGGGGTGACATGATGATTTGACGCCATCAGGGCCGGGTGG 1199
QY 301 TAATATTTACAAATTTGGAGTAACTGATTTGGAACAGGAAAAATTTATCAGAAGATGAAGA 360
DB 1200 TAATATTTACAAATTTGGAGTAACTGATTTGGAACAGGAAAAATTTATCAGAAGATGAAGA 1259
QY 361 AACTGAGTACAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCCTCAGGGTAGC 420
DB 1260 AACTGAGTACAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCCCTCAGGGTAGC 1319
QY 421 TAGCATGGAHAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
DB 1320 TAGCATGGAHAACAAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATC-GAAACT 1378
QY 481 GAAGAGTTGATGATCTGGCTAACAAACACAGAGAAAGACAGGAAATGGAGGAGA 540
DB 1379 GAAGAGTTGATGATCTGGCTAACAAACACAGAGAAAGACAGGAAATGGAGGAGA 1438
QY 541 GCCTCTTGACCTGATCTTTGAAGACTTAAAGCCCAAGTACAAACAACTAAGAGTCTTCA 600
DB 1439 GCCTCTTGACCTGATCTTTGAAGACTTAAAGCCCAAGTACAAACAACTAAGAGTCTTCA 1498
QY 601 AGAAGATCTAGAACAAAGACAAAGTCAAGGTCAATTTCTCTCACTCACTGATGGTGGTAGT 660
DB 1499 AGAAGATCTAGAACAAAGACAAAGTCAAGGTCAATTTCTCTCACTCACTGATGGTGGTAGT 1558
QY 661 TGATGAATCTAGTGGAGTACAGCAACTGCTCTTTGGAAGAACAACTTAAAGGTATTGGG 720
DB 1559 TGATGAATCTAGTGGAGTACAGCAACTGCTCTTTGGAAGAACAACTTAAAGGTATTGGG 1618
QY 721 AGATCATGGGCAACATCTGTAGATGGAGAGAGACCGCTGGGTCTTTTACAGAGCAG 780
DB 1619 AGATCATGGGCAACATCTGTAGATGGAGAGAGACCGCTGGGTCTTTTACAGAGCAG 1678
QY 781 TTCTGACCACTGGAGCGCTGCACCTTCTCTGAGGAACTTCTGGGTGCTGCTACAGCT 840
```

```
DB 1679 CCTTCTCAAATGGCAAGCTCTTACTGAAGAAACAGTGCCTTTTACTGTCATGGCTTTTCAGA 1738
QY 841 GAAAGATGATGAATTAAGCCGCGCAGGACCTATTGAGAGCGACTTTCCAGCAGTTTCAGAA 900
DB 1739 AAAAGAGATGCAATGAACAAAGATTCAACAACTGGCTTTTAAAGATCAAAATGAATGTT 1798
QY 901 GCAGAACGATGTACATAGGGCTTCAAGAGGGGAATTGAAAACCTAAAGAACCTGTATCAT 960
DB 1799 ATCAAGTCTTCAAATAACTGGCCGTTTAAAGCGGATCTAGAAAAGAAAAGCAATCCAT 1858
QY 961 GAGTACTCTTGAGACTGTACGAATATTCTGACAGCAGCCTTTGGAGGACTAGAGAA 1020
DB 1859 GGGCAAACTGTA-----TTCACTCAAAACAAAGATCTTCTTTCAACACTGAG 1904
QY 1021 ACTCTACCAGAGCCACAGAGACTGCCTCTGAGGAGAGAGCCCAAGATGTCACTCGGT 1080
DB 1905 AATAAGTCACTGACCCAGAGAGCGGAAGCATGGCTGGA----- 1942
QY 1081 TCTACGAAGCAGGCTGAGGAGGTCAATCTGAGTGGGAAAAATTTGAACCTGCATCCGC 1140
DB 1943 -----TAACTTTGCCCGGTGTTGGGATAATTAGTCCAAAAAATTTGAA----- 1985
QY 1141 TGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACCTTCAAGAGGCCAC 1200
DB 1986 -----AAGAGTACAGCACAGACCTTGAAGACTCCAGGAACTTCAAGAGGCCAC 2035
QY 1201 GGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGAATCAAGGATCTCTGGCAGCCGT 1260
DB 2036 GGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGTGAATCAAGGATCTCTGGCAGCCGT 2095
QY 1261 GGGCGATCTCTCATTGACTCTCCAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 1320
DB 2096 GGGCGATCTCTCATTGACTCTCTCCAGATCACCTCGAGAAAGTCAAGGCACTTCGAGG 2155
QY 1321 AGAAATTTGGCCCTCTGAAAGAGAACGTGAGCCACGCTCAATGACCTTGTGCGCACCTTAC 1380
DB 2156 AGAAATTTGGCCCTCTGAAAGAGAACGTGAGCCACGCTCAATGACCTTGTGCGCACCTTAC 2215
QY 1381 CACTTTGGCGATTCAGCTCTCACCGTATAAAGCTCTGGAAGACCTTGAAGACCTTGAACACAG 1440
DB 2216 CACTTTGGCGATTCAGCTCTCACCGTATAAAGCTCTGGAAGACCTTGAAGACCTTGAACACAG 2275
QY 1441 ATGGAAGCTTCTGCGAGTGGCCGTGAGGACCGAGTCAAGGCACTTGCATGAAGCCCAAG 1500
DB 2276 ATGGAAGCTTCTGCGAGTGGCCGTGAGGACCGAGTCAAGGCACTTGCATGAAGCCCAAG 2335
QY 1501 GGAATTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTGTCCAGGGTCCCTGGAGAG 1560
DB 2336 GGAATTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTGTCCAGGGTCCCTGGAGAG 2395
QY 1561 AGCCATCTCGCAAAACAAAGTGCCTTACTATATCAACACAGAGACTCAAAACAACTTGTGTG 1620
DB 2396 AGCCATCTCGCAAAACAAAGTGCCTTACTATATCAACACAGAGACTCAAAACAACTTGTGTG 2455
QY 1621 GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGTGAACCTGAAATATGTAGATT 1680
DB 2456 GGACCATCCCAAAATGACAGAGCTCTACAGTCTTTAGTGAACCTGAAATATGTAGATT 2515
QY 1681 CTCAGCTTATAGAGCTGCGCACTGAACTCCGAAAGTCTGCAAGAGGCGCTTTGCTTGGATCT 1740
DB 2516 CTCAGCTTATAGAGCTGCGCACTGAACTCCGAAAGTCTGCAAGAGGCGCTTTGCTTGGATCT 2575
QY 1741 CTTGAGCCCTGTGAGCTGCAATGTATGCTTTGGAACCAAGCAACAACTTCAAGCAAAATGACCA 1800
DB 2576 CTTGAGCCCTGTGAGCTGCAATGTATGCTTTGGAACCAAGCAACAACTTCAAGCAAAATGACCA 2635
QY 1801 GCCCATGGATATCTGACAGATTTAATGTTTGGACCTATTTATGACCGCTCGAGCA 1860
DB 2636 GCCCATGGATATCTGACAGATTTAATGTTTGGACCTATTTATGACCGCTCGAGCA 2695
QY 1861 AGAGCAACAAATTTGGTCAACGCTCCCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCT 1920
DB 2696 AGAGCAACAAATTTGGTCAACGCTCCCTCTCTGCGTGGATATGTGTCTGAACTGGCTGCT 2755
```


QY 1363 CTTGCTCGCAGCTTACACACTTTGGGCAATCAGCTCTCACCGTATAACCTCAGCACTCT 1422
Db 2599 CTTGCTCGCAGCTTACACACTTTGGGCAATCAGCTCTCACCGTATAACCTCAGCACTCT 2658
QY 1423 GGAAGACCTGAACACACAGATGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGCA 1482
Db 2659 GGAAGACCTGAACACACAGATGAAGCTTCTGAGGTGGCGCTCGAGGACCGAGTCAGGCA 2718
QY 1483 GTGCGATGAAGCCCAACAGGACCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTCTGT 1542
Db 2719 GTGCGATGAAGCCCAACAGGACCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTCTGT 2778
QY 1543 CAGGGGTCCCTGGGAGAGAGCCATCTCGCAACAAAGTCCCTACTATATCAACCCAGCA 1602
Db 2779 CAGGGGTCCCTGGGAGAGAGCCATCTCGCAACAAAGTCCCTACTATATCAACCCAGCA 2838
QY 1603 GACTCAAAACAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACAGCTTTTACGTGA 1662
Db 2839 GACTCAAAACAACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACAGCTTTTACGTGA 2898
QY 1663 CTTGATATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGACTGCAGAA 1722
Db 2899 CTTGATATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGACTGCAGAA 2958
QY 1723 GGGCCCTTTTGGATCTCTTGAGCCCTGTCAGCTGTCATGTCATGTCCTTGACCCAGCAAA 1782
Db 2959 GGGCCCTTTTGGATCTCTTGAGCCCTGTCAGCTGTCATGTCATGTCCTTGACCCAGCAAA 3018
QY 1783 CTTCAAGCAAAATGACAGCCCATGATATCTCGAGATTAATTTGTTGACCACTAT 1842
Db 3019 CTTCAAGCAAAATGACAGCCCATGATATCTCGAGATTAATTTGTTGACCACTAT 3078
QY 1843 TTATGACCCCTGGAGCAAGACCAAAATTTGGTCAACGTCCTCTCTGCGTGGATAT 1902
Db 3079 TTATGACCCCTGGAGCAAGACCAAAATTTGGTCAACGTCCTCTCTGCGTGGATAT 3138
QY 1903 GTGCTGAATGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGCTCT 1962
Db 3139 GTGCTGAATGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGCTCT 3198
QY 1963 GTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTTGGAAAGCAAGTACAGATA 2022
Db 3199 GTCTTTTAAACCTGGCATCATTTCCCTGTGTAAGCACATTTTGGAAAGCAAGTACAGATA 3258
QY 2023 CTTTTCAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCTCT 2082
Db 3259 CTTTTCAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGGCTGGGCTCTCT 3318
QY 2083 TCTGCGATGATCTTATCCAA 2101
Db 3319 TCTGCGATGATCTTATCCAA 3337

RESULT 13
US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No.: US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: US-06968
; CURRENT APPLICATION NUMBER: US/10/149, 736
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 5417

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-39

Query Match 76.7%; Score 1611.6; DB 13; Length 5417;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1872; Conservative 0; Mismatches 229; Indels 93; Gaps 3;

QY 1 GAGCTATGCTACACACAGGCTGCTTATGTACCACTCTGTGAGGAGTTCATTTGGCAGTTCATTTGATGAGAG 120
Db 1099 GAGCTATGCTACACACAGGCTGCTTATGTACCACTCTGTGAGGAGTTCATTTGGCAGTTCATTTGATGAGAG 1158
QY 61 TCCTTCACACGATTTGGAGCTCTCTGAGCAAGTTCATTTGGCAGTTCATTTGATGAGAG 120
Db 1159 TCCTTCACACGATTTGGAGCTCTCTGAGCAAGTTCATTTGGCAGTTCATTTGATGAGAG 1218
QY 121 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGATTAATTCGTGGCTTTCTTTC 180
Db 1219 TGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGATTAATTCGTGGCTTTCTTTC 1278
QY 181 TGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCTTATGATGTCGGAAGTTCGTGAAGA 240
Db 1279 TGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCTTATGATGTCGGAAGTTCGTGAAGA 1338
QY 241 CCAGTTCATACCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 300
Db 1339 CCAGTTCATACCTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGGCCGGTTGG 1398
QY 301 TAATATTTCAATTTGGGAAGTAAAGTCTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 360
Db 1399 TAATATTTCAATTTGGGAAGTAAAGTCTGATTTGGAACAGGAAATTTATCAGAGATGAAGA 1458
QY 361 AACTGGAAGTCAAGAGCAGATGATCTCTCTAAATTCAGATGGAATGCTCAGGGTAGC 420
Db 1459 AACTGGAAGTCAAGAGCAGATGATCTCTCTAAATTCAGATGGAATGCTCAGGGTAGC 1518
QY 421 TAGCATGGAAAAACAAGCAATTTACATGAGTTCATGATTTTAAATGGAATCTCCAGATCAGAACT 480
Db 1519 TAGCATGGAAAAACAAGCAATTTACATGAGTTCATGATTTTAAATGGAATCTCCAGATCAGAACT 1578
QY 481 GAAAGAGTTCGAATGATCTGGCTTACAAAAACAGAGAAAGAAACAAGGAAATGAGGAAAGA 540
Db 1579 GAAAGAGTTCGAATGATCTGGCTTACAAAAACAGAGAAAGAAACAAGGAAATGAGGAAAGA 1638
QY 541 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACAACATAAGTGTTCA 600
Db 1639 GCCTCTTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTACAAACAACATAAGTGTTCA 1698
QY 601 AGAAGATCTAGAACCAAGACAGTCAAGGTCAATCTCTCCTCCTCAGATGCTGGTGGTAGT 660
Db 1699 AGAAGATCTAGAACCAAGACAGTCAAGGTCAATCTCTCCTCCTCAGATGCTGGTGGTAGT 1758
QY 661 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAAAGAACAACTTAAAGTATTGGG 720
Db 1759 TGATGAATCTAGTGGAGATCACGCAACTGCTGTTTGGAAAGAACAACTTAAAGTATTGGG 1818
QY 721 AGATCGATGGGCAACATCTGTAGATGGACAGACCCCTGGGTTCTTTTACAGACAG 780
Db 1819 AGATCGATGGGCAACATCTGTAGATGGACAGACCCCTGGGTTCTTTTACAGACAT 1878
QY 781 TTCTGACCACTGGAAGCTGCTGCACTCTTCTCTGACGGAACCTTCTGCTGCTGGTGGTAGT 840
Db 1879 CTTTCTCAATGCGACGCTTACTAGAGACAGTGCCTTTTATGTCATGCTGGCTTTTCA 1938
QY 841 GAAAGATGATGAATTAAGCGCGGAGGACCTATTTGAGGCGGACTTTCCAGAGTTCAGAA 900
Db 1939 AAAAGAGATGCGAGTGAACCAAGATTCACACACTGGCTTTTAAAGATCAAAATGAATGT 1998
QY 901 GCAGACGATGTACATAGGGCC--TTCAAGAGGGAATTTGAAACTTAAAGAACTCTGTAATC 958
Db 1999 ATCAAGTCTTCAAAAACCTGGCCCTTTTAAAGCGGATCTAGAAAGAAAGCAATCCAT 2058

QY 959 ATGAGTACTCTTGAGACTGTACGAATATTT-----CTGACAGAGCAGCGCTTTGGA 1008
Db 2059 GGGCAAACTGTATTAATCACTCAAAAGAGATCTTTCAACACACTGAAGAAATAGTCACTGAC 2118
QY 1009 AGGACTAGAGAACTCTACAGAGAGCCAGAGAGTGCCTCTGAGAGAGAGAGCCAGAA 1068
Db 2119 CCGAAGACGGAAGCATGGCTGTGATAACTTTGCGCGGTGTGGGATAATTTAGTCCAAA 2178
QY 1069 TGTCACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATAC----- 1110
Db 2179 ACTTGAAGAGGTACAGCAGATTTACAGGCTGTACCAACCACTCAGCCATCACTAAC 2238
QY 1111 ----- 1110
Db 2239 ACAGACAACCTGTAATGGAAACAGTAACCTACGCTGACCAACAAGGGAACAGATCCTGGTAAA 2298
QY 1111 ---TGAGTGGGAAAAATTGAACCTGCACTCCGCTGACTGCAGAGAAAAATAGATGAGAC 1167
Db 2299 GCATGCTCAAGAGNACTTCCACACACCTCCCAAAGAGAGCGCATTAAGTGTGA 2358
QY 1168 CTTGAAAGACTCCAGAACTTCAAGAGGCCACGATGAGCTGAGCTCAAGCTGCGCCA 1227
Db 2359 TCTTGAAGACTCCAGAACTTCAAGAGGCCACGATGAGCTGAGCTCAAGCTGCGCCA 2418
QY 1228 AGCTGAGTGTATCAAGGATCTCTGGCAGCCGCTGGGCGATCTCTCAATGACTCTCTCA 1287
Db 2419 AGCTGAGTGTATCAAGGATCTCTGGCAGCCGCTGGGCGATCTCTCAATGACTCTCTCA 2478
QY 1288 AGATCACTCTCGAAGAACTCAAGGCACTTTCAGAGAGAAATTTGGCCCTCTCAAGAGAACT 1347
Db 2479 AGATCACTCTCGAAGAACTCAAGGCACTTTCAGAGAGAAATTTGGCCCTCTCAAGAGAACT 2538
QY 1348 GAGCCAGTCAATGAGCTTCTGCGCAGCTTACCACTTTGGGCACTTCACTCTCAAGTAA 1407
Db 2539 GAGCCAGTCAATGAGCTTCTGCGCAGCTTACCACTTTGGGCACTTCACTCTCAAGTAA 2598
QY 1408 TAACCTCAGCACTCTGGAAGACCTGAAACCAAGATGGAAGCTTCTGAGGCTGGCGTGA 1467
Db 2599 TAACCTCAGCACTCTGGAAGACCTGAAACCAAGATGGAAGCTTCTGAGGCTGGCGTGA 2658
QY 1468 GACCCAGTCAAGGAGTGTATGAAGCCACAGGACTTTGTCAGCACTCTCAGCACTT 1527
Db 2659 GACCCAGTCAAGGAGTGTATGAAGCCACAGGACTTTGTCAGCACTCTCAGCACTT 2718
QY 1528 TCTTTCAGCTCTGTCAGGCTCTCGGAGAGAGCCATCTCGCCAAACAAAGTCCCTA 1587
Db 2719 TCTTTCAGCTCTGTCAGGCTCTCGGAGAGAGCCATCTCGCCAAACAAAGTCCCTA 2778
QY 1588 CTATATCAACACAGAGACTCAAAACACTTGTCTGGGACCACTCCCAAAATGACAGACTCTA 1647
Db 2779 CTATATCAACACAGAGACTCAAAACACTTGTCTGGGACCACTCCCAAAATGACAGACTCTA 2838
QY 1648 CAGTCTTTAGCTGACTGAATATGTAGATCTCAGCTTATAGACTGCGCATGAAGT 1707
Db 2839 CAGTCTTTAGCTGACTGAATATGTAGATCTCAGCTTATAGACTGCGCATGAAGT 2898
QY 1708 CCGAAGACTGCAAGAGCCCTTTGCTTGTATCTTTGAGCCCTGTCAGCTGCAATGTATGC 1767
Db 2899 CCGAAGACTGCAAGAGCCCTTTGCTTGTATCTTTGAGCCCTGTCAGCTGCAATGTATGC 2958
QY 1768 CTTGGACAGCAACCTCAAGCAAAATGACAGCCCATGATGATATCTTGCAGATTTAA 1827
Db 2959 CTTGGACAGCAACCTCAAGCAAAATGACAGCCCATGATGATATCTTGCAGATTTAA 3018
QY 1828 TTGTTTGAACCACTATTATGACCGCTGAGAGAGAGCAACAAATTTGGTCAAGCTCC 1887
Db 3019 TTGTTTGAACCACTATTATGACCGCTGAGAGAGAGCAACAAATTTGGTCAAGCTCC 3078
QY 1888 TCTCTCGTGGATATGTGTGACTGAGCTGCTGATGATTTTATGATACGGGACGAGCAGG 1947
Db 3079 TCTCTCGTGGATATGTGTGACTGAGCTGCTGATGATTTTATGATACGGGACGAGCAGG 3138

QY 1948 GAGGATCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGA 2007
Db 3139 GAGGATCGTGTCTGTCTTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGA 3198
QY 2008 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAG 2067
Db 3199 AGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCAAG 3258
QY 2068 CAGGCTGGGCTCTCTGTGATGATTTCTATCCAA 2101
Db 3259 CAGGCTGGGCTCTCTGTGATGATTTCTATCCAA 3292
RESULT 14
US-09-845-416-6
; Sequence 6, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-6
Query Match 76.2%; Score 1602; DB 13; Length 3999;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;
QY 1 GAGTATGCTTACACACAGGCTGCTTATGTGACACCTCTGACCCCTACACGGAGCCCAATT 60
Db 900 GAGTATGCTTACACACAGGCTGCTTATGTGACACCTCTGACCCCTACACGGAGCCCAATT 959
QY 61 TCCCTTCACAGCACTTTGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTTCAATGATGGAGAG 120
Db 960 TCCCTTCACAGCACTTTGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTTCAATGATGGAGAG 1019
QY 121 TGAAGTAACTGACCGTTATCAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 180
Db 1020 TGAAGTAACTGACCGTTATCAACAGCTTTAGAAAGATATTATCGTGGCTTCTTTC 1079
QY 181 TGCTGAGGACACATTTGAAGCAAGAGAGATTTCTTAATGATGTGAAGTGTGAAGA 240
Db 1080 TGCTGAGGACACATTTGAAGCAAGAGAGATTTCTTAATGATGTGAAGTGTGAAGA 1139
QY 241 CCAAGTTTCACTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGGTGG 300
Db 1140 CCAAGTTTCACTCATGAGGGGTACATGATGGAATTTGACAGCCCATCAGGGCCGGGTGG 1199
QY 301 TAATATTTCAATTTGGGAGTAAAGCTTATGGAACAGGAAATTTATCAGAGATGAAGA 360
Db 1200 TAATATTTCAATTTGGGAGTAAAGCTTATGGAACAGGAAATTTATCAGAGATGAAGA 1259
QY 361 AACTGAAGTCAAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 420
Db 1260 AACTGAAGTCAAGAGCAGATGAATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 1319
QY 421 TAGCATGGAAAAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 480
Db 1320 TAGCATGGAAAAAACAAGCAATTTACATAGAGTTTAAATGATCTCCAGATCAGAACT 1379
QY 481 GAAAGAGTTCAATGACTGGCTTAAACAAAACAGAAAGAAACAAAGGAAATTTGAGGAGGA 540
Db 1380 GAAAGAGTTCAATGACTGGCTTAAACAAAACAGAAAGAAACAAAGGAAATTTGAGGAGGA 1439

541 GCCTCTTGACCTGATCTTTGAAGACCTAAAGCGCAAGTACAAACAATAAGGTGCTTCA 600
1440 GCCTCTTGACCTGATCTTTGAAGACCTAAAGCGCAAGTACAAACAATAAGGTGCTTCA 1499
601 AGAAGATCTAGAACAAAGAAACAAGTCAAGGTCAATTTCTCTCACTCAATGTTGGT 660
1500 AGAAGATCTAGAACAAAGAAACAAGTCAAGGTCAATTTCTCTCACTCAATGTTGGT 1559
661 TGATGAATCTAGTGGAGATCAGCAACTGCTGTTGGAGAACAACTTAAGGTATTGGG 720
1560 TGATGAATCTAGTGGAGATCAGCAACTGCTGTTGGAGAACAACTTAAGGTATTGGG 1619
721 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAGA --- 777
1620 AGATCGATGGCAAAACATCTGTAGATGGACAGAACCGCTGGGTCTTTTACAAGACCA 1679
778 ----- 777
1680 GCGTGAACCTGCTGGAGTCAACCACTATTGGAGCCCTCTCTACTCAAGCTGTACTCT 1739
778 ----- 777
1740 GGTGACAACTGTTGTTACTAAGGAACCTCCATCTCCAACTAGAAATGCCATCTTC 1799
778 ----- 777
1800 CTTGATGTTGGAGTACCTACTCATAGATTACTGCAACAGTTCCCTCGGACCTGGAAAA 1859
778 ----- 777
1860 GTTCTTGCTGGCTTACAGAACTGAAACAACTGCCAATGTCTTACAGGATGTACCCG 1919
778 ----- 777
1920 TAAGAAAGGCTCTTAGAAGACTCCHAGGGAGTAAAGAGCTGATGAAACAATGGCAAGA 1979
778 ----- 777
1980 CTTCAAGGTGAATGAAGCTCACAGATGTTTATCAACCTGGATGAACACAGCCA 2039
778 ----- 777
2040 AAAAACTCCTGAGATCCTTGAAGGTTCCGATGATGAGTCTGTTTACAAGAGCTTTGGA 2099
778 ----- 777
2100 TAAATGAACCTTCAAGTGGAGTGAACCTTCGAAAAAGTCTCTCAACATTAGGTCCCAAT 2159
778 -----CAGTTCTGACAGTGAAGGCTCTGCACCTTTCTCTGAGGAACTTCTGGTGTG 831
2160 GGAAGCCAGTCTGACCACTGGAGGCTCTGCACCTTTCTGAGGAACTTCTGGTGTG 2219
832 GCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCGACTTTCCAGC 891
2220 GCTACAGCTGAAGATGATGAATTAAGCCGGCAGGCACCTATTGAGGCGACTTTCCAGC 2279
892 AGTTCAAGACAGACGATCTACATAGGCTCTCAGAGGGAATTGAACCTTAAGAAC 951
2280 AGTTCAAGACAGACGATCTACATAGGCTCTCAGAGGGAATTGAACCTTAAGAAC 2339
952 TGTAACTCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCTTTTGAAGG 1011
2340 TGTAACTCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAGCTTTTGAAGG 2399
1012 ACTAGAGAACTTACAGAGGCCAGAGGCTGCTCTGAGGAGAGGCCAGATGT 1071
2400 ACTAGAGAACTTACAGAGGCCAGAGGCTGCTCTGAGGAGAGGCCAGATGT 2459
1072 CACTCGGCTTCTACGAAGCAGGCTCAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT 1131
2460 CACTCGGCTTCTACGAAGCAGGCTCAGGAGGTCAATACTGAGTGGGAAAAATTGAACCT 2519

1132 GCATCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGAACTTCA 1191
2520 GCATCCGCTGACTGCGCAGAGAAAAATAGATGAGACCCCTTGAAGACTCCAGAACTTCA 2579
1192 AGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCCCTG 1251
2580 AGAGGCCACGATGAGCTGGACCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCCCTG 2639
1252 GCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCCGAAAAAGTCAAGGC 1311
2640 GCAGCCGCTGGGCGATCTCTCATTTGACTCTCTCCAAAGATCACTCCGAAAAAGTCAAGGC 2699
1312 ACTTCAGAGAGAAATTCGCCCTCTGAAAGAGAAAGTGAAGCCACGTCATGACTCTGCTCG 1371
2700 ACTTCAGAGAGAAATTCGCCCTCTGAAAGAGAAAGTGAAGCCACGTCATGACTCTGCTCG 2759
1372 CCAGCTTACCACTTTGGGCAATTCAGCTCTCACGTTAACTTCACTCAGCACTCTGGAAGACT 1431
2760 CCAGCTTACCACTTTGGGCAATTCAGCTCTCACGTTAACTTCACTCAGCACTCTGGAAGACT 2819
1432 GAAACACAGATGGAAGCTTCTGAGGTGCCCTGAGGACCGAGTCAAGCACTGAGCTGATGA 1491
2820 GAAACACAGATGGAAGCTTCTGAGGTGCCCTGAGGACCGAGTCAAGCACTGAGCTGATGA 2879
1492 AGCCACAGAGGACTTTCGAGCATCTCAGCACTTCTTCCAGCTGCTCCAGGCTCC 1551
2880 AGCCACAGAGGACTTTCGAGCATCTCAGCACTTCTTCCAGCTGCTCCAGGCTCC 2939
1552 CTGGGAGAGGACCATCTCGCCAAAACAAAGTGCCTACTATATCAACACGAGACTCAAAAC 1611
2940 CTGGGAGAGGACCATCTCGCCAAAACAAAGTGCCTACTATATCAACACGAGACTCAAAAC 2999
1612 AACTTCGCTGGACCATCTCCAAATGACAGACTCTACAGTCTTTAGCTGACTGAATAA 1671
3000 AACTTCGCTGGACCATCTCCAAATGACAGACTCTACAGTCTTTAGCTGACTGAATAA 3059
1672 TGTCAAGTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGCTGCAAGGCCCTTTG 1731
3060 TGTCAAGTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGCTGCAAGGCCCTTTG 3119
1732 CTTGGATCTCTGAGCCTGTGAGCTGATGATGCTTGGACCGACCACTCAAGCA 1791
3120 CTTGGATCTCTGAGCCTGTGAGCTGATGATGCTTGGACCGACCACTCAAGCA 3179
1792 AATGACCGACCGCATGATCTGAGATTAATGCTGAGACTTAACTTTGACCACTATTATGACCG 1851
3180 AATGACCGACCGCATGATCTGAGATTAATGCTGAGACTTAACTTTGACCACTATTATGACCG 3239
1852 CTTGGACGACGACCAACAATTTGCTCAAGCTTCCCTCTCTCGGTGGATGCTGTGAA 1911
3240 CTTGGACGACGACCAACAATTTGCTCAAGCTTCCCTCTCTCGGTGGATGCTGTGAA 3299
1912 CTTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGCTCTCTTTAA 1971
3300 CTTGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGCTCTCTTTAA 3359
1972 AACTGGCATCATTTCCCTGTGTAAGACCAATTTGGAAGCAAGTACAGATACCTTTCAA 2031
3360 AACTGGCATCATTTCCCTGTGTAAGACCAATTTGGAAGCAAGTACAGATACCTTTCAA 3419
2032 GCAAGTGGCAAGTTCACAGGATTTGTGACGCGAGCTGGGCTCTCTTCTGATGA 2091
3420 GCAAGTGGCAAGTTCACAGGATTTGTGACGCGAGCTGGGCTCTCTTCTGATGA 3479
2092 TTCTATCCAA 2101
3480 TTCTATCCAA 3489

RESULT 15
US-09-845-416-28
; Sequence 28, Application US/09845416
; Publication No. US20030171312A1

; GENERAL INFORMATION:

; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 4966
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-845-416-28

Query Match 76.2%; Score 1602; DB 13; Length 4966;
Best Local Similarity 81.1%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 0; Indels 489; Gaps 1;

Qy 1 GAGTATGCTACACACAGCGTCTTATGTCACACACTCTGACCCCTACCGAGCCCATTT 60
Db 1657 GAGTATGCTACACACAGCGTCTTATGTCACACACTCTGACCCCTACCGAGCCCATTT 1716
Qy 61 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGAGAG 120
Db 1717 TCCTTCACAGCATTTGGAAGCTCTGAAGACAAAGTCAATTTGGCAGTTCAATGATGAGAG 1776
Qy 121 TGAAGTAAACCTGCGACCGTTATCAACACAGCTTTAGAGAGTATTCGTGGCTTCTTTC 180
Db 1777 TGAAGTAAACCTGCGACCGTTATCAACACAGCTTTAGAGAGTATTCGTGGCTTCTTTC 1836
Qy 181 TSCGTGAGGACACATTGCAACACAGAGAGATTTCTAATGATGCGAAGTGTGAAAGA 240
Db 1837 TSCGTGAGGACACATTGCAACACAGAGAGATTTCTAATGATGCGAAGTGTGAAAGA 1896
Qy 241 CCAGTTTCACTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGGTGG 300
Db 1897 CCAGTTTCACTACTCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGCGGGGTGG 1956
Qy 301 TAATATTCTCAATTTGGGAAGTAGTGTATGTTGGAACAGAGAAATATCAGAAATGAAGA 360
Db 1957 TAATATTCTCAATTTGGGAAGTAGTGTATGTTGGAACAGAGAAATATCAGAAATGAAGA 2016
Qy 361 AACTGGAAGTCAAGAGCAGATGAATCTCTTAATTCAGATGCGAATGCTCAGGCTAGC 420
Db 2017 AACTGGAAGTCAAGAGCAGATGAATCTCTCTTAATTCAGATGCGAATGCTCAGGCTAGC 2076
Qy 421 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTATTTAATGGATCTCCAGAATCAGAACT 480
Db 2077 TAGCATGGAAGAAACAAAGCAATTTACATAGAGTATTTAATGGATCTCCAGAATCAGAACT 2136
Qy 481 GAAAGAGTTGAATGACTGGCTACAAACACAGAAAGAAAGAAACAGGAAATGGAGGAAGA 540
Db 2137 GAAAGAGTTGAATGACTGGCTACAAACACAGAAAGAAAGAAACAGGAAATGGAGGAAGA 2196
Qy 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTACAAACATAAAGTGCTTCA 600
Db 2197 GCCTCTTGGACCTGATCTTGAAGACCTTAAACCCCAAGTACAAACATAAAGTGCTTCA 2256
Qy 601 AGAAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTCACTCACAATGGTGGTAGT 660
Db 2257 AGAAGATCTAGAACAGAAACAAGTCAGGGTCAATTTCTCACTCACAATGGTGGTAGT 2316
Qy 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAACACTTAAAGGTATGGG 720
Db 2317 TGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAGAACACTTAAAGGTATGGG 2376
Qy 721 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGA --- 777
Db 2377 AGATCGATGGGCAACATCTGTAGATGGACAGAGACCGCTGGGTTCTTTTACAAGACCA 2436

Qy 778 ----- 777
Db 2437 GCCTGACCTAGCTCTCTGACTGACCACTATTGGAGCCTCTCTTACTCAGACTGTACTCT 2496
Qy 778 ----- 777
Db 2497 GGTGACACAACCTGTGTTACTTAAGAAACTGCGCATCTCCAAACTAGAAATGCCATCTTC 2556
Qy 778 ----- 777
Db 2557 CTTGATGTTGGAGGTACCTACTCATAGATTACTGCAACAGTTTCCCTGGACCTGGAATA 2616
Qy 778 ----- 777
Db 2617 GTTCTTTCCTGCTTACAGAAAGCTGAAACAACTGCAATGCTCTACAGGATGCTTACCG 2676
Qy 778 ----- 777
Db 2677 TAAGAAAGGCTCTTAGNAGACTCCAAGGAGTAAAGAGCTGATGAACAATGGCAAGA 2736
Qy 778 ----- 777
Db 2737 CCTCAAGGTGAAATTTGAAGCTCACACAGATGTTTATCAACAACCTGGATGAAACACAGCCA 2796
Qy 778 ----- 777
Db 2797 AAAAATCCTGAGATCCCTGGAGGTTCCGATGATGCGAGTCTCTGTACAAGACGTTTGA 2856
Qy 778 ----- 777
Db 2857 TAACATGAATTTCAAGTGGAGTGAATTCGGAAAAAGTCTCTCAACATTAGTCCCATTT 2916
Qy 778 ----- 777
Db 2917 GGAAGCCAGTTTCTACAGTGAAGAGCTCTGCACTTCTCTGAGGAACTTCTGCTGTG 2976
Qy 832 GCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTATTGGAGCGCACTTTCAGC 891
Db 2977 GCTACAGCTGAAAGATGATGAATTAAGCCGCGAGCACCTATTGGAGCGCACTTTCAGC 3036
Qy 892 AGTTGAGAGCAGAAACAGATGTATAGGGCTTCAAGAGGGAATTTGAAACTAAGAAC 951
Db 3037 AGTTGAGAGCAGAAACAGATGTATAGGGCTTCAAGAGGGAATTTGAAACTAAGAAC 3096
Qy 952 TGTAAATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAGG 1011
Db 3097 TGTAAATCATAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCTTTGGAGG 3156
Qy 1012 ACTAGAGAACTCTTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGT 1071
Db 3157 ACTAGAGAACTCTTACAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCCAAGATGT 3216
Qy 1072 CACTCGGCTTCTAGAAAGCAGCTGAGGAGTCAATCTGAGTGGGAAAAATTTGACCT 1131
Db 3217 CACTCGGCTTCTAGAAAGCAGCTGAGGAGTCAATCTGAGTGGGAAAAATTTGACCT 3276
Qy 1132 GCATCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAATTTCA 1191
Db 3277 GCATCTCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTCCAGGAATTTCA 3336
Qy 1192 AGAGCCACGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCAAGGATCTCTG 1251
Db 3337 AGAGCCACGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATCAAGGATCTCTG 3396
Qy 1252 GCAGCCGTTGGGCTCTCTCTTATGATCTCTTCAAGATCACTCGAGAAAGTCAAGGC 1311
Db 3397 GCAGCCGTTGGGCTCTCTCTTATGATCTCTTCAAGATCACTCGAGAAAGTCAAGGC 3456
Qy 1312 ACTTCGAGGAAATTCGCTCTGAAAGAGAAAGCTGAGCCACGTCATGACCTTGTCTG 1371
Db 3457 ACTTCGAGGAAATTCGCTCTGAAAGAGAAAGCTGAGCCACGTCATGACCTTGTCTG 3516
Qy 1372 CCAGCTTACCACTTTGGGCAATTCAGCTCTCCAGCTATACTTACCTCAGCACTCTGGAAGACCT 1431

```
Db      |||||CCAGCTTACCACTTTGGGCACTCCTCAGCTATACCTCAGCACTCTGGAGACCT 3576
Qy      |||||GAACACAGATGAGAGCTTCTCAGAGTGGCCGTCAGAGCCGAGTCAGGACGCTGCATGA 1491
Db      |||||GAACACAGATGAGAGCTTCTCAGAGTGGCCGTCAGAGCCGAGTCAGGACGCTGCATGA 3636
Qy      |||||AGCCACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCC 1551
Db      |||||AGCCACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCC 3696
Qy      |||||CTGGGAGAGAGCCATCTGCCAAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAAC 1611
Db      |||||CTGGGAGAGAGCCATCTGCCAAAACAAAGTGCCCTACTATATCAACACAGAGACTCAAAAC 3756
Qy      |||||AACTTCTGGGACCACTCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATAA 1671
Db      |||||AACTTCTGGGACCACTCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACCTGAATAA 3816
Qy      |||||TGTCAAGTTCTCAGCTTATAGGACTGCCATGAACCTCCGAGACTGCAAGAGGCCCTTTG 1731
Db      |||||TGTCAAGTTCTCAGCTTATAGGACTGCCATGAACCTCCGAGACTGCAAGAGGCCCTTTG 3876
Qy      |||||CTTGGATCTCTTGAGCCTGTACGCTGCTATGATGCTTGGACGAGCAACCTCAAGCA 1791
Db      |||||CTTGGATCTCTTGAGCCTGTACGCTGCTATGATGCTTGGACGAGCAACCTCAAGCA 3936
Qy      |||||AAATGACAGCCCATGGATATCTCGAGATTATTAATTGTTTGACCACCTATTATGACCG 1851
Db      |||||AAATGACAGCCCATGGATATCTCGAGATTATTAATTGTTTGACCACCTATTATGACCG 3996
Qy      |||||CCTGGAGCAAGAGCAACAATTTGTTGTAACGCTCCTCTCGGTGGATATGTCTGAA 1911
Db      |||||CCTGGAGCAAGAGCAACAATTTGTTGTAACGCTCCTCTCGGTGGATATGTCTGAA 4056
Qy      |||||CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCGCTGCTCTTTTAA 1971
Db      |||||CTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGATCCGCTGCTCTTTTAA 4116
Qy      |||||AACTGATCATTTCCCTGTGTAAAGCAATTTGGAGACAAGTACAGATACCTTTTCAA 2031
Db      |||||AACTGATCATTTCCCTGTGTAAAGCAATTTGGAGACAAGTACAGATACCTTTTCAA 4176
Qy      |||||GCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCCTTCTGCATGA 2091
Db      |||||GCAAGTGGCAAGTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCCTTCTGCATGA 4236
Qy      |||||TTCTATCCAA 2101
Db      |||||TTCTATCCAA 4246
```

Search completed: February 2, 2004, 17:39:28
Job time : 720.863 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:04:35 ; Search time 3939.6 Seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-12_COPY_900_3000

Perfect score: 2101

Sequence: 1 gagctatgctacacacagg.....ttctgcatgtattctatccaa 2101

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estm:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estc:*
10: gb_estc:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pig:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	787	37.5	3870	11 BC036103	BC036103 Homo sapi
2	629.8	30.0	3056	11 AK044536	AK044536 Mus muscu
3	514.6	24.5	1490	11 BC009242	BC009242 Homo sapi
4	509.2	24.2	3753	11 AK081426	AK081426 Mus muscu

C	5	509	24.2	595	14	CB177816
	6	505	24.0	824	9	AL556247
	7	503.4	24.0	801	14	CB981394
	8	503.4	24.0	1098	13	EX355572
	9	494	23.5	620	13	BQ640063
	10	494	23.5	797	14	CB960722
	11	480	22.8	728	14	CB228986
	12	471.4	22.4	770	10	BG719710
	13	455.8	21.7	591	2	HSM083521
	14	453.8	21.6	652	10	BB629984
	15	453.8	21.6	1298	11	AK087829
	16	453.8	21.6	2135	11	AK013510
	17	453.8	21.6	4437	11	AK036936
	18	452.2	21.5	1384	11	AK075809
	19	449	21.4	704	10	BB610411
	20	439	20.9	578	2	HSM075761
	21	428.6	20.4	854	9	AL196693
	22	423	20.1	665	14	BY742604
	23	410.4	19.5	599	10	BB666688
	24	407.2	19.4	554	14	CB613696
	25	404.6	19.3	3051	11	BC036095
	26	404.4	19.2	493	14	CA888041
	27	398.2	19.0	717	14	CB527785
	28	397.4	18.9	495	14	CA894775
	29	392.2	18.7	650	14	BY714491
	30	388.6	18.5	644	13	BU313510
	31	383.4	18.2	483	14	CA893902
	32	381.8	18.2	423	9	AA460476
	33	380.2	18.1	684	9	AL641565
	34	375.4	17.9	514	9	AL871560
	35	374.8	17.8	898	10	BF182055
	36	369.6	17.6	843	14	CA988247
	37	368.6	17.5	633	9	AL796733
	38	349.6	16.6	546	4	EX516597
	39	334.6	15.9	794	13	BU424348
	40	332.6	15.8	541	9	AL894729
	41	330	15.7	608	9	AL672616
	42	329	15.7	532	10	BE334408
	43	326.2	15.5	655	14	CA558919
	44	325.6	15.5	532	12	BU036615
C	45	324.6	15.4	402	14	CB547284

ALIGNMENTS

```

RESULT 1
BC036103
LOCUS BC036103 3870 bp mRNA linear HTC 04-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:5274415, mRNA.
ACCESSION BC036103
VERSION BC036103.1 GI:23271310
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
Contact: MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305

```

REMARK COMMENT

Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 48 Row: f Column: 7
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
 1..3870
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5274415"
 /tissue_type="Brain, hippocampus"
 /clone_lib="NIH MGC 95"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

BASE COUNT 1321 a 677 c 763 g 1109 t

ORIGIN

Query Match 37.5%; Score 787; DB 11; Length 3870;
 Best Local Similarity 94.8%; Pred. No. 5.3e-154;
 Matches 814; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

1 GAGCTATGCTACACACAGCGTCTATCTCACCACCTGACCTACACGAGCCCAT 60
 1043 GAGCTATGCTACACAGCGTCTATCTCACCACCTGACCTACACGAGCCCAT 1102
 61 TCCTTCACAGCATTGGAAGCTCCTGAAGACAGTCAATTGGCAGTTCATTGATGAGAG 120
 1103 TCCTTCACAGCATTGGAAGCTCCTGAAGACAGTCAATTGGCAGTTCATTGATGAGAG 1162
 121 TGAAGTAAACCTGACCGTTCATCAACAGCTTTAGAAGAGTATTATCGTGCTTCTTTC 180
 1163 TGAAGTAAACCTGACCGTTCATCAACAGCTTTAGAAGAGTATTATCGTGCTTCTTTC 1222
 181 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 240
 1223 TGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAGA 1282
 241 CCAGTTTCATCTCATGAGGGGTACATGATGATGATGATGATGATGATGATGATGATGATG 300
 1283 CCAGTTTCATCTCATGAGGGGTACATGATGATGATGATGATGATGATGATGATGATGATG 1342
 301 TAATATTCTCAATTTGGGAGTAACTGATGATGATGATGATGATGATGATGATGATGATGATG 360
 1343 TAATATTCTCAATTTGGGAGTAACTGATGATGATGATGATGATGATGATGATGATGATGATG 1402
 361 AACTGAAGTACAAGACAGATGAATCTCTTAAATCAAGATGGGAATGCTCAGGGTAGC 420
 1403 AACTGAAGTACAAGACAGATGAATCTCTTAAATCAAGATGGGAATGCTCAGGGTAGC 1462
 421 TAGCATGGAATAAACAAGCAATTTATAGAGTTTAAATGATCTCCAGATCAGAACT 480
 1463 TAGCATGGAATAAACAAGCAATTTATAGAGTTTAAATGATCTCCAGATCAGAACT 1522
 481 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAAACAGAGAAATGGAGGAAGA 540
 1523 GAAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGAAACAGAGAAATGGAGGAAGA 1582
 541 GCCTCTTGAGCCTGATCTTGAAGCCTTAAACCGCAAGTACACACATAAGTGTCTTCA 600
 1583 GCCTCTTGAGCCTGATCTTGAAGCCTTAAACCGCAAGTACACACATAAGTGTCTTCA 1642
 601 AGAAGATCTAGAACAAACAAGTCAGGGTCAATTTCTCTCAGTCACATGATGATGATGATG 660
 1643 AGAAGATCTAGAACAAACAAGTCAGGGTCAATTTCTCTCAGTCACATGATGATGATGATG 1702

661 TGATGAATCTAGTGGAGATCAACCAACTGCTGTTTGGGAAGCAACACTTAAGTATGGG 720
 1703 TGATGAATCTAGTGGAGATCAACCAACTGCTGTTTGGGAAGCAACACTTAAGTATGGG 1762
 721 AGATCGATGGGCAACATCTGTAGATGGACAGACCCGCTGGTCTTTTCAACAGACAG 780
 1763 AGATCGATGGGCAACATCTGTAGATGGACAGACCCGCTGGTCTTTTCAACAGACAT 1822
 781 TTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGACAGAACTTCTGCTGGTCAGACT 840
 1823 CCTTCTCAANTGCAACGCTTACTGAGAACAGTCCCTTTTAGTGATGCTGCTTTCAGA 1882
 841 GAAAGATGATGAATTAAGC 859
 1883 AAAAGAAGATGCAGTGAAC 1901

RESULT 2
 AK044536 3056 bp mRNA linear HTC 05-DEC-2002
 LOCUS Mus musculus adult retina cDNA, RIKEN full-length enriched library,
 DEFINITION clone:A930019F21 product:dystrophin, muscular dystrophy, full
 insert sequence.
 ACCESSION AK044536
 VERSION AK044536.1 GI:26090404
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636

2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159

3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Hono, H., Kasukawa, T., Saito, R.,
 Kado, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischnann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
 Quackenbush, J., Schiraldi, L. M., Staubli, F., Suzuki, R., Tomita, M.,
 Weng, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, C., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyehaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 3056)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Retina RNA was provided by Dr. Stefano Gustinich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES

Location/Qualifiers

1..3056

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:A930019F21"

/db_xref="taxon:10090"

/clone="A930019F21"

/tissue_type="retina"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

misc_feature 1..3056

/note="dystrophin, muscular dystrophy (MGD | MGI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)"

BASE COUNT 1072 a 567 c 607 g 810 t

ORIGIN

Query Match 30.0%; Score 629.8; DB 11; Length 3056;

Best Local Similarity 83.4%; Pred. No. 3.8e-121;

Matches 715; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 1 GAGCTATGCTTACACAGCGTCTTATGTCACACCTTGACCTTACAGGAGCCCATTT 60

944 GAGTATGCTTACACAGCGTCTTATGTCACACCTTGACCTTACAGGAGCCCATTT 1003

Db

QY 61 TCCTTCCACAGCATTTGAGCTCTTGAACACAGTCATTTGGCAGTTTCATTCATGGAGAG 120

Db 1004 TCCTTCCACAGCATTTGAGCTCTTGAAGCTCCACAGAGACAAGTCATTCGACATTCATTGATGAGAGAC 1063

QY 121 TGAAGTAAACCTTGGACCGTTATCAACACAGCTTTAGAGAAGTATTATCGTGGCTTCTTTTC 180

Db 1064 GGAAGTAAATCTGGTAGTTACCAACTGCTTTAGAGAAGTACTTTTCATGCTTCTTTTC 1123

QY 181 TCGTAGGACACATTTGCAAGCAACAGGAGAGATTTCTTAATGATGTGGAGAGTGGTGAAGA 240

Db 1124 TCGCAGGAGTACATTTGCGAGCAACAGGAGAGATTTCAAATGATGTGAAGAAAGTGAAGA 1183

QY 241 CCAAGTTTCTATCTCATGAGGGGTGATGATGATGATTTGACAGCCCATCAGGCGCGGTTGG 300

Db 1184 ACAGTTTCAATGCTCATGAGGGATTCATGATGATTCGACATCTCATCAAGGACTTGTGG 1243

QY 301 TAAATTTCTACAAATTTGGGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 360

Db 1244 TAAATTTCTACAAATTTAGGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1303

QY 361 AACTGAGTACAGAGCAGCAGTAAATCTCTTAATTCAGATGGGAATGCTCAGGCTAGC 420

Db 1304 AGCTGAAGTGAAGCAACAATGAATCTCTTAATTCAGATGGGAATGCTCAGGCTAGC 1363

QY 421 TAGCATGGAAAAACCAAGCAATTTACATAGAGTTTTTAATGGATCTCCAGAAATCAGAAACT 480

Db 1364 TAGCATGGAAAAACCAAGCAATTTACCAAAAGTTCTTAATGGATCTCCAGAAATCAGAAACT 1423

QY 481 GAAAGAGTGAATGATGCTGCTTAAACAAAACAGAGAAACAGAGAAACAGAGAAACAGAGAAAC 540

Db 1424 AAAAGAACTAGATGATGCTGCTTAAACAAAACAGAGAAACAGAGAAACAGAGAAACAGAGAAAC 1483

QY 541 GCCTCTTGGACCTGATCTTGAAGACCTTAAACAAAACAGAGAAACAGAGAAACAGAGAAACAGAGAAAC 600

Db 1484 GCCTCTTGGACCTGATCTTGAAGATCTTAAATGCCAAGTACAAACATAGGCTTCA 1543

QY 601 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCACTCACTCACTCACTCACTCACTCACTCACT 560

Db 1544 AGAAGATCTAGAACAAAGCAAGTCAAGGTCAATTTCTCACTCACTCACTCACTCACTCACTCACTCACT 1603

QY 661 TGATGAATCTAGTGGAGATCAGCAACTGCTGTTTGGAGAACCAACTTAAAGTATTGGG 720

Db 1604 TGATGAATCAGCGGTGATGATGCAACAGCTGTTTGGAGAACCAACTTAAAGTATTGGG 1663

QY 721 AGATCGATGGCAACATCTGATAGTGGACAGACCGCTGGTGTCTTTTCAAGACAG 780

Db 1664 AGATCGATGGCAATATCTGATAGTGGACAGACCGCTGGTGTCTTTTCAAGACAT 1723

QY 781 TTCTGACCACTGCAAGCGTCTGCACCTTTCTCTGACGAACTTCTGCTGCTGCTGCTGCTGCTGCTGCT 840

Db 1724 TCTTCTAAATGGCAGCATTTTACTGAGAACAGTGCCTTTTATGATGATGCTTTCAGA 1783

QY 841 GAAAGATGATGAATTA 857

Db 1784 AAAAGAGATGCAATCA 1800

RESULT 3

BC009242

LOCUS

DEFINITION

Homo sapiens, dystrophin (muscular dystrophy, Duchenne and Becker types), includes DXS142, DXS164, DXS206, DXS230, DXS239, DXS268, DXS269, DXS270, DXS272, clone IMAGE:3029414, mRNA.

BC009242

VERSION

BC009242.1 GI:14714379

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1490)

Strausberg, R.

Direct Submission

Submitted (06-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nhl.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahney, Brin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 10 Row: j Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5032284
This clone has the following problem: retained intron.

FEATURES
source
1..1490
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3029414"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC 17"
/lab_host="DH10B-R"
/note="Vector: pOTB7" 358 t

BASE COUNT 505 a 299 c 328 g 358 t

ORIGIN

Query Match. 24.5%; Score 514.6; DB 11; Length 1490;
Best Local Similarity 99.2%; Pred. No. 4.2e-97;
Matches 517; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAGTATGCTTACACAGAGCTGCTTATGTCACACAGCTTGCACACAGTCTGACCTTACACGAGGCCATT 60
DB 969 GAGTATGCTTACACAGAGCTGCTTATGTCACACAGCTTGCACACAGTCTGACCTTACACGAGGCCATT 1028

QY 61 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAGTCAATTTGGCAGTTCATTGATGGAGAG 120
DB 1029 TCCTTCACAGCATTTGGAAGCTCCTGAAGACAGTCAATTTGGCAGTTCATTGATGGAGAG 1088

QY 121 TGAAGTAACTGACGCGTTATCAACAGCTTTAGAGAGTATATCGTCTCTTTC 180
DB 1089 TGAAGTAACTGACGCGTTATCAACAGCTTTAGAGAGTATATCGTCTCTTTC 1148

QY 181 TGCTGAGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA 240
DB 1149 TGCTGAGACACATTGCAAGCACAAGAGAGATTTCTAATGATGTGGAAGTGTGGAAGA 1208

QY 241 CCAGTTTCATCTCATGAGGGGTACATGATGATGATTTGACAGCCATCAGGCGCGGGTTGG 300
DB 1209 CCAGTTTCATCTCATGAGGGGTACATGATGATGATTTGACAGCCATCAGGCGCGGGTTGG 1268

QY 301 TAATATTCTCAATTGGGAAGTGAAGTGTGGAACAGGAATATTCAGAGATCAAGA 360
DB 1269 TAATATTCTCAATTGGGAAGTGAAGTGTGGAACAGGAATATTCAGAGATCAAGA 1328

QY 361 AACTGAAGTCAAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCTCAGGTTAGC 420
DB 1329 AACTGAAGTCAAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCTCAGGTTAGC 1388

QY 421 TAGCATGGAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGATCAGAACT 480
DB 1389 TAGCATGGAACAAAGCAATTTACATAGATTTTAAATGGATCTCCAGATCAGAACT 1448

QY 481 GAAAGAGTTGAATGACTGCGGTAAACAAAAACAGAGAAAGAA 521
DB 1449 GAAAGAGTTGAATGACTGCGGTAAACAAAAACAGAGAAAGAA 1489

RESULT 4
AK081426
LOCUS
DEFINITION

AK081426 3753 bp mRNA linear HTC 05-DEC-2002
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:Cl30016k19 product:cystrophin related protein 2, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK081426.1 GI:26349154
HTC; CAP trapper.
Mus musculus (house mouse)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
9279253
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Washiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Washiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, K., Iizawa, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barah, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, C., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohseki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3753)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Onno, M., Ohsato, N., Okazaki, Y., Saico, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

Location/Qualifiers
1. .3753
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB.C130016K19"
/db_xref="taxon:10090"
/clone="C130016K19"
/tissue_type="head"
/dev_stage="16 days embryo"
181..3054
/note="unnamed protein product; dystrophin related protein match=807"
2 (MGD|GI:107432, GI|U43520, evidence: BLASTN, 99%,
putative)
/codon_start=1
/protein_id="BAC38217.1"
/db_xref="GI:26349155"
/translation="MQPLVWQCPYTLPRCEHWAADRFHSHSLLENTCPQPVRAAY
TIPAPPDGGDFCLSPKLVGTGATGPLEPSAMNLCWNEIKKSHNLRARLEAPSD
LSGLQLPRLIIDWLSQDEELSAQLFGQDVALVQOEKETHAAPMEVYKSGPYIS
SVLESQAFLSQHPFELESQESKXTSPRIQNLRFVWKQATVASSELWEKLTAR
CVDQRIHIEHTLHLEIQAMBELSTLQAEGRVATWEPIDGLFIDSLPEHIQAIK
LFKEEFPVKGVLNDLAHQLAIGDVLHLSMENSRLBOINRWKLOVSVVAERLQK
LOADHDFGPGSOHFLSTSVQVPERAFSPNKVPYVNHQAOTCWDHPKMTLYQTL
ADLNLIKFSAYRTAMKLRVQKALDLVLTLTALSFNEHDLQASEHWDVVEVIHC
LTALYERLEERGLVNPVCDMSLNWLNLFDSGRGNRLSPKTIAGLCGTEV
KEKQLYLFVSANGSCDQRIHALLHEAIQVPRQGEVAAFSGNSVPSVRSFRF
STGKPIEASQFLEWNLEPQVHVAHLVRVIAEQVHQTKSCIQCPCKGFRYR
SLKQFNVDICQTLFTGRASGNKLHPIMEYVPTTNSNMRDFAITLKNKFRSQY
PSKHPOGYLPVOSVLESCESTPASSPMLPHADTHSRIEHPSRLAEVMSQCFEN
DSLSPDSIDEDAYLLRHSSPIIDRPAPQQAQPCSWATESGKELILAHLEDNRIL
LQELRLKQWHEAEAPFLVGSRAEPDHNRELLAEKILRHQKSLRLETRQWIL
EDHNQLSEQLRLQPPDESNGSGLASSPSQSGSHRSHRKGQTTTPDTE
VADDVSGSQDVSLEIDIMEKLHRAFPVSRSSDVTANTILLAS"
3731..3736
/note="putative"
3753
/note="putative"
BASE COUNT 927 a 1021 c 938 g 867 t
ORIGIN

CDS

Query Match 24.2%; Score 509.2; DB 11; Length 3753;
Best Local Similarity 61.1%; Pred. No. 6.3e-96;
Matches 823; Conservative 0; Mismatches 523; Indels 0; Gaps 0;

polyA_signal

polyA_site
BASE COUNT 927 a 1021 c 938 g 867 t
ORIGIN

756 ACCGCTGGGTTCTTTTACAAGACAGCTTCTGACCACTGGAAGCGTCTGCACCTTTCTCTGC 815
DB
455 ACCTCGCGCTCGCTAGAGGCTTCTCAGACCTCAGTGGAAAACCTCAGCTCCCGCTCC 514
DB
816 AGGAACCTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCGCGGAGGACCTATTG 875
DB
515 GAGAGATTATTGACCTGGCTCAGCCAAAAGGATGAGGATTTGTCACTCAGCTCCCTTGC 574
DB
876 GAGGCGACTTTCCAGCAGTTCAAGACGACAGATGATACATAGGCGCTTCAAGAGCGAAT 935
DB
575 AAGGGGATGTGCGCTCGGTACACAGGAGAGGAGACACATCGACCTTCTATGGAAGAAG 634
DB
936 TGAAGACTAAGAACTGTATATCATGACTCTTTGAGACTGTAGCAATATTTCTGACAG 995
DB
635 TCAAGTCTTAAGGGCGCTTACATCTCTCTGTGCTTGAATCCGCCAGGCTTTTCTGTGCC 694
DB
996 AGCAGCTTTTGAAGGACTAGAGAACTCTACCAGGAGCCAGAGAGCTGCTCTCTGAGG 1055
DB
695 AGCATCCATTTGAAGANTTAGAGAGTCTCAATCTGAGAGTAAGATACCTCCCCAGAC 754
DB
1056 AGAGAGCCAGAAATCTCACTCGGCTTTACGAAAGCAGGCTGAGAGTCAATATCTAGT 1115
DB
755 AGCGGATTCNAACCTTAGTCGCTTTGATGAAAGCAGGCAACAGTGCCTGAGCTGT 814
DB
1116 GCGAAAATTTGAACCTGCACTCCGCTGACTGCGAGAGAAATATAGTAGACCTTTGAAA 1175
DB
815 GGGAGAAGCTGACAGCGCTGTGTAGATCAGCATCGCCACATTTAGAGCATACTCTGGAAC 874
DB
1176 GACTCCAGAACTTCAAGAGGCGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1235
DB
875 ATCTATTGGAGATCCAAAGGGGCAATGGAGGAATTTGAGCAGTACTTTGACCAAGCAGG 934
DB
1236 TGATCAAGGATCTCTGGCAGCGCTGGCGGATCTCTCATTTGACTCTCTCAAGATCACC 1295
DB
935 GAGTCCGAGCAGCTGGGAGCCATAGGAGATCTCTTTATTTGATTCCTCCAGAGCATA 994
DB
1296 TCGAGAAAGTCAAGGCACTTCGAGAGAGAAATTCGCGCTCTGAAAGAGAACTGAGCCACG 1355
DB
995 TCCAGCCATCAAGCTATTCAAGAGAGATTTCTCTCTGTGAAAGATGGGTGAAGTAG 1054
DB
1356 TCAATGACCTTGTGCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCA 1415
DB
1055 TGAATGATCTGGCGCCACCACTTGCCTCTCTGATGTGCACTTGTCAATGAGAAATTTCCA 1114
DB
1416 GACTCTCGAAGACCTGAAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1475
DB
1115 GGGCTCTGGAAACAGATCAACATCCGGTGGAAACAGCTCCAGGTGTGAGTGTGAGAGGC 1174
DB
1476 TCAGGAGCTGCATGAAGCCCAAGGCACTTTGGTCCAGATCTCAGCACTTTCTTTTCCA 1535
DB
1175 TTAGCAACTCCAGATGCGCCACCGGACCTTTGGGCTGGGTCAAGCACTTCTCTCTCCA 1234
DB
1536 CGTGTCTCAGGTTCCCTGGGAGAGGCACTCTCGCCAAACAAAGTGCCTTACTATATCA 1595
DB
1235 CTTCTGTCCAAGTTCCCTGGGAAAGAGCATTTTCTCCCAATAAAGTCCCTTACTACATCA 1294
DB
1596 ACCAGAGACTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTTACCAAGTCTT 1655
DB
1295 ACCAGAGCTCAGACCACTGCTGGGACCATCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1354
DB
1656 TAGCTGACCTGAATTAATGTTCAGTTTCAGCTTATAGGATGCGCATGAAATCCGAGAGAC 1715
DB
1355 TAGCGATCTGAACAAACATTAAGTTCTCAGCTTACCAGCTGCCATGAAGTCCGAGAG 1414
DB
1716 TCGAAGGCGCTTTGCTTGGATCTCTTGAGCGCTGTCAGCTGTCATGTGATGCTTGAACC 1775
DB
1415 TCCAGAGGCGCTTGGCGCTGGATCTGGTAATTTAACTACAGCTCTCGAGATCTTCAATG 1474
DB
1776 AGCAACACTCAAGCAAAATGACCAAGCCCATGGATATCTCTGCAAGTATTAATTTGTTGA 1835
DB
1475 AGCATGACTTCAGGCGCCAGTGAACATGTGATGATGTGGTGGAGGTCAATTCACCTGCTTGA 1534

1836 CCACTATTATGACCGCTGGAGCAAGAGACACAAATTTGGTCAACGTCCTCTCTCGC 1895
 1535 CTGCTTGTATGAACGACTGGAGGAGAAAGAGGATCTCTGTCAATGTGCCACTGTGTG 1594
 1896 TGGATATGTCTGAACCTGGCTGCTGAATGTTATGATACGGAGCAAGAGGAGATCC 1955
 1595 TAGACATGAGCCTCAACTGCTCTCAATGTTTGTATAGTGTGCGAGTGAAGAATGC 1654
 1956 GTCTCTCTGCTTTTAAACATGGCATCATTTCCCTGTGTAAAGACACATTTGGGAAGCAAGT 2015
 1655 GAGCAATGCTTTTAAAGATGGCATCGCATGCTGTGTGCGAGGAGTGAAGAAAC 1714
 2016 ACAGATACCTTTCAAGCAAGTGGGAAGTTCAACAGGATTTGTACACGAGCCAGGCTGG 2075
 1715 TTCAGTATCTCTTCAGCCAAAGTAGCAATTCAGGCAGCCAGTGTGATCAACGCCATCTCG 1774
 2076 GCCTCTCTGTCATGATTTCTATCCAA 2101
 1775 GTGCCCTGCTTCATGAGCAATTCAA 1800

RESULT 5
 CBI177816/c
 LOCUS
 DEFINITION is21c01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6553129 3', similar to SW:DMD_HUMAN P11532 DYSTROPHIN. [1] ; mRNA sequence.
 ACCESSION CBI177816
 VERSION CBI177816
 KEYWORDS EST.
 SOURCE CBI177816.1 GI:28186206

ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 595)
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y., and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished
 Other ESTs: is21c01.y1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@bioh.harvard.edu
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@wustl.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 448.
 Location/Qualifiers

FEATURES
 source
 1..595
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6553129"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NcoI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St.

Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."
 BASE COUNT 146 a 136 c 116 g 197 t
 ORIGIN

Query Match 24.2%; Score 509; DB 14; Length 595;
 Best Local Similarity 100.0%; Pred. No. 5.7e-96;
 Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGTATGCGCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACAGGAGCCCAAT 60
 Db 509 GAGTATGCGCTACACACAGGCTGCTTATGTCAACACCTCTGACCCCTACAGGAGCCCAAT 450
 QY 61 TCCTTCAAGCATTTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTCATTGATGAGAG 120
 Db 449 TCCTTCAAGCATTTGGAAGCTCTCTGAAGCAAGTCAATTTGGCAGTTCATTGATGAGAG 390
 QY 121 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTTC 180
 Db 389 TGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAGTATTATCGTGGCTTCTTTTC 330
 QY 181 TGTGAGGACATTTGCAAGCAAGAGAGATTTCTAATGATGTGGAAAGTGTGTAAGA 240
 Db 329 TGTGAGGACATTTGCAAGCAAGAGAGATTTCTAATGATGTGGAAAGTGTGTAAGA 270
 QY 241 CCAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGGGTTGG 300
 Db 269 CCAGTTTCACTCATGAGGGGTACATGATGATTTGACAGCCCATCAGGCGGGTTGG 210
 QY 301 TAATATTCTACAATTGGGAAGTAAAGCTGATTTGAAACAGGAAATTTATCAAGATGAAGA 360
 Db 209 TAATATTCTACAATTGGGAAGTAAAGCTGATTTGAAACAGGAAATTTATCAAGATGAAGA 150
 QY 361 AACTGAATCAAGAGCAGATGATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 420
 Db 149 AACTGAATCAAGAGCAGATGATCTCTTAATTTCAAGATGGGAATGCTCAGGGTAGC 90
 QY 421 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 480
 Db 89 TAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATGGATCTCCAGAAATCAGAAACT 30
 QY 481 GAAAGAGTTGAATGACTGGCTAACAAAA 509
 Db 29 GAAAGAGTTGAATGACTGGCTAACAAAA 1

RESULT 6
 AL556247
 LOCUS
 DEFINITION AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens cDNA clone CSODK001YB17 5-PRIME, mRNA sequence.
 ACCESSION AL556247
 VERSION AL556247.2 GI:31278051
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 824)
 Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 15, 2001 this sequence version replaced gi:12898746.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1955.r
 Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK001CA09Q1.
 Location/Qualifiers

FEATURES


```

RESULT 8
BX365572      1098 bp      mRNA      linear      EST 05-MAY-2003
LOCUS      BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION      CDNA clone CS0DK001YB21 5-PRIME, mRNA sequence.
ACCESSION      BX365572
VERSION      BX365572.1 GI:30366927
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1098)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished
COMMENT      Contact: Genoscope
Genoscope, Centre National de Sequencage
BP 191 9106 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSLAK001ZB11QPL.
Location/Qualifiers
1..1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB21"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      286 a 271 c 268 g 262 t 11 others
ORIGIN
Query Match      24.0%; Score 503.4; DB 13; Length 1098;
Best Local Similarity 99.8%; Pred. No. 9e-95;
Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1597 CCACGAGACTCAAAACAACTGCTGGGACCATCCCAAAATGACAGAGCTTACAGCTCTTT 1656
DB      204 CCACGAGACTCAAAACAACTGCTGGGACCATCCCAAAATGACAGAGCTTACAGCTCTTT 263

QY      1657 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAGACT 1716
DB      264 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACTCCGAGACT 323

QY      1717 GCAGAGGCCCTTTGCTTGATCTCTTGAGCCTGTGTCAGCTGATGATGATGCTTGAGCA 1776
DB      324 GCAGAGGCCCTTTGCTTGATCTCTTGAGCCTGTGTCAGCTGATGATGATGCTTGAGCA 383

QY      1777 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCTTCGAGATTATTAATTTGTTGAC 1836
DB      384 GCACAACCTCAAGCAAAATGACAGCCCATGGATATCTTCGAGATTATTAATTTGTTGAC 443

QY      1837 CACTATTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCGT 1896
DB      444 CACTATTATGACCGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCGT 503

QY      1897 GGATATGTCCTGACGCTGCTGATGTTATGATAGGGACGACAGGAGGATCCG 1956
DB      504 GGATATGTCCTGACGCTGCTGATGTTATGATAGGGACGACAGGAGGATCCG 563

QY      1957 TGTCTCTGCTTTTAAACTGGCATCAATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTA 2016
DB      564 TGTCTCTGCTTTTAAACTGGCATCAATTTCCCTGTGTAAAGCACATTTTGGAGACAAGTA 623

QY      2017 CAGATACCTTTTCAAGCAAGTGCAGATTCAACAGGATTTTGTGACCGAGGCTGGG 2076
DB      624 CAGATACCTTTTCAAGCAAGTGCAGATTCAACAGGATTTTGTGACCGAGGCTGGG 683

QY      2077 CTTCTCTTCTGCATGATCTATCCAA 2101
DB      684 CTTCTCTTCTGCATGATCTATCCAA 708

RESULT 9
BX640063      620 bp      mRNA      linear      EST 15-JUL-2002
LOCUS      BX640063 Homo sapiens Human Retina cDNA (Un-normalized, unamplified): hd/he
DEFINITION      Homo sapiens cDNA clone he23g04 5', mRNA sequence.
ACCESSION      BX640063
VERSION      BX640063.1 GI:21764522
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 620)
AUTHORS      Wistow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A., Touchman
J.W., Bouffard, G., Smith, D. and Peterson, K.
TITLE      Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
JOURNAL      Contact: Wistow G
COMMENT      Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 23 row: g column: 04
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="he23g04"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retina cDNA (Un-normalized, unamplified)"
/notes="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the SuperScript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-pGACTGTTCTAGATCGAGCGCGCCCTT]15-3'.
EST analysis was performed on the unamplified library
at the NIH Intramural Sequencing Center (NISC)."
BASE COUNT      165 a 148 c 144 g 163 t
ORIGIN
Query Match      23.5%; Score 494; DB 13; Length 620;
Best Local Similarity 100.0%; Pred. No. 7.8e-93;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1608 AAACAATTGCTGGGACCATCCCAAAATGACAGAGCTTACAGCTCTTTAGCTGACCTGA 1667
DB      1 AAACAATTGCTGGGACCATCCCAAAATGACAGAGCTTACAGCTCTTTAGCTGACCTGA 60

```

```
QY 1568 ATATGTCAGATCTCAGCTTATAGGACTGCCATGAACCTCCGAGACTGCAGAGGCC 1727
D 11
D 61 ATATGTCAGATCTCAGCTTATAGGACTGCCATGAACCTCCGAGACTGCAGAGGCC 120
QY 1728 TTGCTGGATCTCTTGAGCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1787
D 121 TTGCTGGATCTCTTGAGCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 180
QY 1788 AGCAAAATGACAGCCCATGATGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1847
D 181 AGCAAAATGACAGCCCATGATGATGCTGATGCTGATGCTGATGCTGATGCTGATG 240
QY 1848 ACCGCTGGAGCAGGAGCAGCAAAATTTGTCAGCTGCTGCTGCTGCTGCTGCTGCT 1907
D 241 ACCGCTGGAGCAGGAGCAGCAAAATTTGTCAGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 1908 TGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1967
D 301 TGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 1968 TTAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2027
D 361 TTAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 2028 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGGAGGCTGCTGCTGCT 2087
D 421 TCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGGAGGCTGCTGCTGCT 480
2Y 2088 ATGATTTATCCAA 2101
D 481 ATGATTTATCCAA 494

RESULT 10
LOCUS CB960722
DEFINITION AGENCOURT_13761995 NIH_MGC_147 Homo sapiens cDNA clone
IMAGE:30344479 5', mRNA sequence.
ACCESSION CB960722
VERSION CB960722.1 GI:30216839
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 797)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM380 row: n column: 08
High quality sequence stop: 620.
Location/Qualifiers
1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30344479"
/tissue_type="Human Placenta"
/lab_host="PH108 Tona"
/clone_lib="NIH_MGC_147"
/note="Organ: placenta; Vector: pBluescriptR; Site: 1;
all-XhoI; Site 2: BamH; Oligo-dr primed using primer
5'-TTTTTTTTTTTTTTT-3', size-selected for average
```

```
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGM/NHGRI, National Institutes of Health). Note: This is
a NIH MGC library."
BASE COUNT 199 a 208 c 199 g 191 t
ORIGIN
Query Match 23.5%; Score 494; DB 14; Length 797;
Best Local Similarity 99.8%; Pred. No. 8e-93;
Matches 505; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1597 CCAGAGACTCAAAACAATCTGCTGGGACCATCCAAATAGCAGAGCTCTACAGTCTTT 1656
D 168 CCAGAGACTCAAAACAATCTGCTGGGACCATCCAAATAGCAGAGCTCTACAGTCTTT 227
QY 1657 AGCTGACCTGAATTAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAGACT 1716
D 228 AGCTGACCTGAATTAATGTCAGATTCTCAGCTTATAGGACTGCCATGAACCTCCGAGACT 287
QY 1717 GCAGAGGCCCTTTGCTTGGATCTCTTGGAGCCTGTGAGCTGCTGATGCTGCTGCTGCTG 1776
D 288 GCAGAGGCCCTTTGCTTGGATCTCTTGGAGCCTGTGAGCTGCTGATGCTGCTGCTGCTG 347
QY 1777 GCACAACCTCAAGCAAAATGACGAGCCCATGAGTATCTCAGCTGCTGCTGCTGCTGCTG 1836
D 348 GCACAACCTCAAGCAAAATGACGAGCCCATGAGTATCTCAGCTGCTGCTGCTGCTGCTG 407
QY 1837 CACTATTATGACCGCTGGAGCAAGAGCACAACATTTGGTCAACGCTCCCTCTCTGCTG 1896
D 408 CACTATTATGACCGCTGGAGCAAGAGCACAACATTTGGTCAACGCTCCCTCTCTGCTG 467
QY 1897 GGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1956
D 468 GGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 527
QY 1957 TGTCCTGCTCTTTAAACTGGCATCATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2016
D 528 TGTCCTGCTCTTTAAACTGGCATCATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587
QY 2017 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAGCAGGATTTTGTGACGAGCGAGGCTGG 2076
D 588 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAGCAGGATTTTGTGACGAGCGAGGCTGG 647
QY 2077 -CCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2101
D 648 CCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 673

RESULT 11
LOCUS CB228986
DEFINITION AGENCOURT_1149247 NIHCD_Rh_Ovi Macaca mulatta cDNA clone
IMAGE:6884820 5', mRNA sequence.
ACCESSION CB228986
VERSION CB228986.1 GI:28280564
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 728)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
CDNA Library Preparation: CLONTECH
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```


encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001).

Yamanaka, I., Kiyosawa, H., Kondo, S. T., Saito, T., Shinagawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Azakawa, T., Ishii, Y. and Hayashizaki, Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct. Func. Genomics* 2 pre, 172-186 (2001).

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

FEATURES
· SOURCE

```
Location/Qualifiers
1. 652
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9930028B14"
/sex="female"
/tissue_type="vagina"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN f
vagina"
```

vagen" Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTITTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCGATTATTAATATACCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(-) after bulk excision from Lambda

FLC 1.1	
BASE COUNT	172 a 160 c 148 g 172 t
ORIGIN	
172 a 160 c 148 g 172 t	

```
Query Match      21.6%; Score 453.8; DB 10; Length 652;
Best Local Similarity 93.7%; Pred. No. 2e-84;
Matches 473; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

Qy	1597	CCACGAGACTCAAAACA	CTTGTGTTGGG	ACCATCCCAAAATGACAGAGTCTACAGTCTTT	1656
Db	45	CCACGAGACCCAAACCA	CTTGTGTTGGG	ACCATCCCAAAATGACAGAGTCTACAGTCTTT	104

QY 1657 AGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACT 1716

Db 105 AGCTGACCTGAATAATGTCAGGTTCTCCGGTATAGGACTGCCATGAAGCTCAGAAGGCT 164

QY 1717 GCAGAGGCCCTTGTCTTGGATCTCTTGAGCCTGTCAGCTGCATGTGATGCCTTGGACCA 1776

D_b 165 CCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAAGCTGCATGTGATGCCCTGGACCA 224

1777 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTATTGTTTGAC 1836

Db 225 GCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCGCAGATAATTAACGTGTTGAC 284

Qy 1837 CACTATTATGACCGCCTGGAGCAAGACACAACAATTGGTCAACGTCCCTCTCTCGGT 1896

285 TACAA TTTATGATCGTCTGGAGCAAGAGCACAAATCTGGTCAATGTCCCTCTCTGTGT 344

1897 GGATATGTGTCTGAACTGGCTGCTGAATGTTTATGATACGGACGACAGGGAGGATCCG 1956

DB 345 GGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACGGGACGAACAGGGAGGATCCG 404

QY 195 / TGTCCCTGTCCTTTTAAAC1GGCATCATTTCCCTGTGTAAAGCACATTTGGTAAGACAAGIA 2016

DB 405 TGTCCTGTCCTTTAAACACGGCAATCATTTCTCTGCTGTTAAAGCACACATTGGAAAGACAAAGTA 407

Accession	Gene	Strain	Position	Sequence	Length
QY	2017	CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTTGTGACCAAGGCAGGCTGGG	2076		
Db	465	CAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTTGTGACCAAGGCAGGCTGGG	524		
QY	2077	CTTCCTTCTGCATGATTCATCCAA	2101		
Db	525	TCCTCTTCTGCATGATTCATCCAA	549		

RESULT 15	AK087829	1298 bp	mRNA	linear	HTC 05-DEC-2002
LOCUS	AK087829				
DEFINITION	Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330026B12 product:dystrophin, muscular dystrophy, full insert sequence.				

ACCESSION AK087829
VERSION AK087829.1 GI:26352777
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
1. Carminci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

2
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

Carinici, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
AUTHORS		Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE		20530913
PUBMED		11076861

4
REFERENCE
AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itch, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Iizawa, M., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fletschmann, W., Gaasterland, I., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashina, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Togo-Oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsuke, S. and Hayashizaki, Y.

TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409 (6821), 685-690 (2001)
MEDLINE
21085660

11217851
 5
 PUBLISHED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 6 (bases 1 to 1298)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, W.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gscl.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 CDS library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in RIKEN
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers
 1..1298
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM,DB:E330026B12"
 /db_xref="taxon:10090"
 /clone="E330026B12"
 /sex="female"
 /tissue_type="ovary"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="2 days pregnant adult"
 120..1142
 /notes="unnamed protein product; dystrophin, muscular
 dystrophy (MGD)|G1:94909, GB|NM_007568, evidence: BLASTN,
 100%, match=999)
 putative"
 /codon_start=1
 /protein_id="BAC40019.1"
 /db_xref="GI:26352778"
 /translation="MREHLKGHEQTQTDWHPKWTLYOSLADLNVRFSAYRTAMKL
 RLQKALCLDLLSACDADLNKQNDQMDILQINCLTYDLRQEHNNLVN
 VPLQCDMLNVLNVDTGTGTRVLSFKTGIIISLCKAHLEDKRYLFLKQVASTGP
 CDQRGLLLHDSIQIPRQGEVASFVSRCFQFANNRPETLFAALFLDWMR
 LEFOSMWLVLEVAATFAKHOAKNICKECPPIIGPRYSLKHFNYDQCSCFPFG
 RVAKGHKHPYVEYCTPTTSGSDVRDFAKVLNKFRTKRYFAKHPRMGYLEVQTVLE
 GDNNET"
 BASE COUNT 343 a 332 c 296 g 327 t
 ORIGIN
 Query Match 21.6%; Score 453.8; DB 11; Length 1298;
 Best Local Similarity 93.7%; Pred. No. 2.1e-84;
 Matches 473; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 QY 1597 CCAGGAGCTCAACAACTGCTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTT 1656
 DB 140 CCAGGAGCCCAACCACTTGTGGGACCAACCCCAAAATGACAGAGCTCTACAGTCTTT 199

Search completed: February 2, 2004, 06:03:17
 Job time : 3941.93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:14:55 ; Search time 7128.21 Seconds
(without alignments)
11483.979 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000

Perfect score: 2001
Sequence: 1 ggcagtcattgatggagag.....cattgagccaaagtgtccgga 2001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_man.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1820.8	91.0	5417	6	AX538619	AX538619 Sequence
2	1332	66.6	4402	6	E30220	E30220 Shortened d
3	1269.2	63.4	4402	6	E30219	E30219 Shortened d
4	11254	62.7	5339	6	AX538620	AX538620 Sequence
5	1132.2	56.6	5462	6	AX538621	AX538621 Sequence
6	1002.6	50.1	5952	6	AR304538	AX304538 Sequence
7	1002.6	50.1	5952	6	AX114289	AX114289 Sequence
8	1002.6	50.1	8689	6	AX538622	AX538622 Sequence
9	1002.6	50.1	11443	6	AX538624	AX538624 Sequence
10	1002.6	50.1	12057	6	AX538627	AX538627 Sequence
11	1002.6	50.1	12446	9	HSDMDR	X14298 Human mRNA
12	1002.6	50.1	13957	6	AX409637	AX409637 Sequence
13	1002.6	50.1	13957	6	AX538581	AX538581 Sequence
14	1002.6	50.1	13957	9	HUMDYS	M18533 Homo sapien
15	991.6	49.6	13977	6	AR220819	AR220819 Sequence
16	885.8	44.3	13887	4	APC70485	APC70485 Caris fam
17	863.8	43.5	13815	6	AX306153	AX306153 Sequence
18	869.8	43.5	13815	6	AX538582	AX538582 Sequence
19	869.8	43.5	13907	6	MUSDYSA	M68859 Mouse dystr
20	869.8	43.5	13907	6	AR093392	AR093392 Sequence
21	869.8	43.5	19307	6	AR142592	AR142592 Sequence
22	776.6	38.8	3275	10	MUSDYS	M18025 Mouse dystr
23	735.4	36.8	4075	6	E30221	E30221 Shortened d
24	731.2	36.5	3747	6	E30218	E30218 Shortened d
25	722.6	36.1	13575	5	GGDYS	X13369 Chicken mRN
26	677.6	33.9	3163	6	E30223	E30223 Shortened d
27	615.4	30.8	2654	5	FSCDYSTRO	M37645 Torpedo cal
28	612.2	30.6	630	9	HSDMDF1	X06179 Human fetal
29	569	28.4	2110	9	HUMDMDX	M92650 Human Duch
30	569	28.4	4658	9	BC028720	BC028720 Homo sapi
31	534	26.7	3521	5	AP339031	AP339031 Danio rer
32	528	26.4	1966	9	HSDMDA1	X06178 Human adult
33	518	25.9	1240	10	RNAPD3	G69767 R.norvegicu
34	514.6	25.7	10705	10	RNAJ2967	AJ002967 Rattus no
35	506.6	25.3	3161	10	MMGUTRPH	X83506 M.musculus
36	506.6	25.3	11096	6	AX538584	AX538584 Sequence
37	506.6	25.3	11096	10	MMV12229	Y12229 M.musculus
38	490.6	24.5	6045	6	A63605	A63605 Sequence 7
39	490.6	24.5	6045	6	AR281528	AR281528 Sequence
40	490.6	24.5	6059	6	AX107972	AX107972 Sequence
41	490.6	24.5	10302	6	AX538583	AX538583 Sequence
42	490.6	24.5	10302	9	HSMUPS	X69086 H.sapiens m
43	490.6	24.5	10320	6	A63607	A63607 Sequence 9
44	490.6	24.5	10320	6	AR281529	AR281529 Sequence
45	475	23.7	3073	10	AF195788	AF195788 Rattus no

ALIGNMENTS

RESULT 1	AX538619	AX538619	Sequence 39 from Patent WO0229056.	5417 bp	DNA	linear	PAT 23-NOV-2002
AX538619	LOCUS	AX538619	Sequence 39 from Patent WO0229056.				
	DEFINITION	AX538619	Sequence 39 from Patent WO0229056.				
	ACCESSION	AX538619	Sequence 39 from Patent WO0229056.				
	VERSION	AX538619.1	GI:25271163				
	KEYWORDS						
	SOURCE		synthetic construct				
	ORGANISM		synthetic construct				
			artificial sequences.				
REFERENCE	1						
AUTHORS			Chamberlain,J.S. and Harper,S.Q.				
TITLE			Mini-dystrophin nucleic acid and peptide sequences				
JOURNAL			Patent: WO 0229056-A 39 11-APR-2002.				
			THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)				

FEATURES		Location/Qualifiers
source	1..5417	/organism="synthetic construct"
		/mol_type="genomic DNA"
		/db_xref="taxon:32630"
		/notes="Synthetic"
BASE COUNT	1700 a	1192 c 1182 g 1343 t
ORIGIN		
Query Match 91.0%; Score 1820.8; DB 6; Length 5417;		
Best Local Similarity 92.8%; Pred. No. 0;		
Matches 1999; Conservative 0; Mismatches 2; Indels 157; Gaps 2;		
Qy	1	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA 60
Db	1199	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA 1258
Qy	61	GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGACGACACAGGAGAGATTCTTAAT 120
Db	1259	GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGACGACACAGGAGAGATTCTTAAT 1318
Qy	121	GATGTGGAAGTGTGAAAGACCACTTTCTACTCATGAGGGGTACATGATGATTGACA 180
Db	1319	GAATGTGGAAGTGTGAAAGACCACTTTCTACTCATGAGGGGTACATGATGATTGACA 1378
Qy	181	GCCCATCAGGCGCGGTGGTAATATTTCTACAAATTGGGAAGTAAGCTGATTGGAAACGGA 240
Db	1379	GCCCATCAGGCGCGGTGGTAATATTTCTACAAATTGGGAAGTAAGCTGATTGGAAACGGA 1438
Qy	241	AAATTATCAGACATCAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGA 300
Db	1439	AAATTATCAGACATCAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGA 1498
Qy	301	TGGGAATGCTCAGGTAGCTAGCTAGCAAAACCAAGCAATTTACATAGATTTTAATG 360
Db	1499	TGGGAATGCTCAGGTAGCTAGCTAGCAAAACCAAGCAATTTACATAGATTTTAATG 1558
Qy	361	GATCTCCAGATC-GAACTGAAGAGTTGAATGACTGCTTAACAAACACAGAAAGA 419
Db	1559	GATCTCCAGAAATCAGAAATCGAAGAGTTGAATGACTGCTTAACAAACACAGAAAGA 1618
Qy	420	ACAAGGAAATGAGGAAGAGCGCTCTTGACCTGATCTTTGAAGACTTAAACGCCAAGTA 479
Db	1619	ACAAGGAAATGAGGAAGAGCGCTCTTGACCTGATCTTTGAAGACTTAAACGCCAAGTA 1678
Qy	480	CAACAACATGAAGTGTCTCAAGAGATCTAGAACAGAAACAAGTCAGGTCAATTTCTCTC 539
Db	1679	CAACAACATGAAGTGTCTCAAGAGATCTAGAACAGAAACAAGTCAGGTCAATTTCTCTC 1738
Qy	540	ACTCACATGTTGTTGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 599
Db	1739	ACTCACATGTTGTTGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 1798
Qy	600	GAACAACATTAAGTATGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 659
Db	1799	GAACAACATTAAGTATGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGC 1858
Qy	660	TGGGTTCTTTTACAAGACATCTTCTCAATGGCAACGTTCTTACTGAAGAACAGTGCCTT 719
Db	1859	TGGGTTCTTTTACAAGACATCTTCTCAATGGCAACGTTCTTACTGAAGAACAGTGCCTT 1918
Qy	720	TTTAGTGCATGGCTTTTCAGAAAAGAGATGCAAGTGAACAAGATTTCACAACACTGGCTTT 779
Db	1919	TTTAGTGCATGGCTTTTCAGAAAAGAGATGCAAGTGAACAAGATTTCACAACACTGGCTTT 1978
Qy	780	AAAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGGCGTTTAAAGCGGATCTA 839
Db	1979	AAAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGGCGTTTAAAGCGGATCTA 2038
Qy	840	GAAGAAGAAAAGCAATCCATGGGCAAACTGTATTCTCAACCAAGATCTTCTTTCAACA 899
Db	2039	GAAGAAGAAAAGCAATCCATGGGCAAACTGTATTCTCAACCAAGATCTTCTTTCAACA 2098
Qy	900	CTGAAGATAAAGTCAGTACCCAGAGACGGAAGCATGCTGTGATAACTTTGCCCCTGT 959
Db	2099	CTGAAGATAAAGTCAGTACCCAGAGACGGAAGCATGCTGTGATAACTTTGCCCCTGT 2158
Qy	960	TGGGATAAATTTAGTCCAAAACTTTGAAAAAGATACAGCACAGA----- 1002
Db	2159	TGGGATAAATTTAGTCCAAAACTTTGAAAAAGATACAGCACAGATTTTCACAGGCTGTCAAC 2218
Qy	1003	----- 1002
Db	2219	ACCACTCAGCCATCACTAACACAGACAACTGTAAATGGAAACAGTAACACTAGGTGACCA 2278
Qy	1003	----- 1002
Db	2279	AGGGAACAGATCCTGTGTAAGCATGTCTAAGAGAACTTCCACACCACTCCCTCCCAAG 2338
Qy	1003	-----CCCTTGAAGACTCCAGAACTTCAAGAGGCGCCACGGATGAG 1043
Db	2339	AAGAGGAGATTACTGTGGATCTTTGAAGACTCCAGAACTTCAAGAGGCGCCACGGATGAG 2398
Qy	1044	CTGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGCAT 1103
Db	2399	CTGACCTCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGCAT 2458
Qy	1104	CTCTCATTTGACTCTCTCCAGATCACCCTCGAAGATCACTCGAAGAGTCAAGGCAATTCAGGAGAAAT 1163
Db	2459	CTCTCATTTGACTCTCTCCAGATCACCCTCGAAGATCACTCGAAGAGTCAAGGCAATTCAGGAGAAAT 2518
Qy	1164	GGGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTTGCTCGCCAGCTTACCCTTTG 1223
Db	2519	GGGCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTTGCTCGCCAGCTTACCCTTTG 2578
Qy	1224	GGATTCAGCTCTACCGTATAACCTCAGCACTCTCGAAGACCTGAAACACAGATGGAAG 1283
Db	2579	GGATTCAGCTCTACCGTATAACCTCAGCACTCTCGAAGACCTGAAACACAGATGGAAG 2638
Qy	1284	CTTCTGAGGTGGCGTTCAGGACCGAGTCAGGACGTGATCAAGCCCAACACAGGACTTT 1343
Db	2639	CTTCTGAGGTGGCGTTCAGGACCGAGTCAGGACGTGATCAAGCCCAACACAGGACTTT 2698
Qy	1344	GGTCCAGACTCTCAGCACTTTTCTTCCAGTCTGTCAGGGTCCCTGGGAGAGAGCAATC 1403
Db	2699	GGTCCAGACTCTCAGCACTTTTCTTCCAGTCTGTCAGGGTCCCTGGGAGAGAGCAATC 2758
Qy	1404	TCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAAACTGCTGGGAGCCAT 1463
Db	2759	TCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAAACTGCTGGGAGCCAT 2818
Qy	1464	CCCAAAATGACAGAGCTTACCAAGTCTTTAGCTGACCTGAATAATGTCAAGTTCTCAGCT 1523
Db	2819	CCCAAAATGACAGAGCTTACCAAGTCTTTAGCTGACCTGAATAATGTCAAGTTCTCAGCT 2878
Qy	1524	TATAGGACTGCCATGAACTCCGAAAGACTGCAGAGCCCTTTGCTTGGATCTCTTGAGC 1583
Db	2879	TATAGGACTGCCATGAACTCCGAAAGACTGCAGAGCCCTTTGCTTGGATCTCTTGAGC 2938
Qy	1584	CTGTCACTGCATGTGATGCTTGGACCAAGCAAACTCAAGCAAAATGACACGCCCATG 1643
Db	2939	CTGTCACTGCATGTGATGCTTGGACCAAGCAAACTCAAGCAAAATGACACGCCCATG 2998
Qy	1644	GATATCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGCTTGGAGCAAGAGCAC 1703
Db	2999	GATATCTGCAGATTATTAATTTGTTGACCACTATTTATGACCGCTTGGAGCAAGAGCAC 3058
Qy	1704	AACAATTTGTTCAACGTCCTCTCTGCTGGATATGTCTGAACCTGGCTGCTGGAATGTT 1763
Db	3059	AACAATTTGTTCAACGTCCTCTCTGCTGGATATGTCTGAACCTGGCTGCTGGAATGTT 3118
Qy	1764	TATGATACGGGACGAAACAGGAGGATCCGTGTCCTGCTTTTAAATCTGCAATTTCC 1823
Db	3119	TATGATACGGGACGAAACAGGAGGATCCGTGTCCTGCTTTTAAATCTGCAATTTCC 3178
Qy	1824	CTGTGTAAGCACATTTGGAGAACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA 1883

Db 3179 CTGTGAAGCACATTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA 3238
QY 1884 ACAGGATTTTGTGACAGCGCAGCTGGGCTCTTCTGCGATGATCTATCCAAATTC 1943
Db 3239 ACAGGATTTTGTGACAGCGCAGCTGGGCTCTTCTGCGATGATCTATCCAAATTC 3298
QY 1944 AGACAGTTGGTGAAGTTGTCATCTTTGGGGGAGTAAACATGAGCCAAAGTCCGGA 2001
Db 3299 AGACAGTTGGTGAAGTTGTCATCTTTGGGGGAGTAAACATGAGCCAAAGTCCGGA 3356

RESULT 2
E30220
LOCUS Shortened dystrophin. 4402 bp DNA linear PAT 18-JUN-2001
DEFINITION Shortened dystrophin.
ACCESSION E30220
VERSION E30220.1 GI:13017027
KEYWORDS JP 1999318467-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4402)
AUTHORS Sinichi T.
TITLE Shortened dystrophin
JOURNAL Patent: JP 1999318467-A 3 24-NOV-1999;
COMMENT SCIENCE & TECH AGENCY NATIONAL CENTER OF NEUROLOGY AND PSYCHIATRY
OS Unidentified
PN JP 1999318467-A/3
PD 24-NOV-1999
PF 08-MAY-1998 JP 1998142134
PR
PI SINICHI TAKEDA
PC C12N15/09,A61K48/00,C12N15/00
CC Strandedness: Both;
CC Topology: Linear;
FH Key
FT source
FT Location/Qualifiers
/organism='Unidentified'.
1. 4402
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 1339 a 984 c 1010 g 1069 t
ORIGIN

Query Match 66.6%; Score 1332; DB 6; Length 4402;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1677; Conservative 0; Mismatches 0; Indels 325; Gaps 2;

QY 1 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1195 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1254
QY 61 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 120
Db 1255 GTATTATCGTGGCTCTTTCTGCTGAGGACACATTGCAAGCACAGGAGAGATTTCTAAT 1314
QY 121 GATGTGAAGTGTGAAGAACCAAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 180
Db 1315 GATGTGAAGTGTGAAGAACCAAGTTTCATCTCATGAGGGGTACATGATGATTTGACA 1374
QY 181 GCCCATCAGGGCCGGTGTGTAATTTCTACATTTGGGAAGTAAAGTGAATTCGAAACAGGA 240
Db 1375 GCCCATCAGGGCCGGTGTGTAATTTCTACATTTGGGAAGTAAAGTGAATTCGAAACAGGA 1434
QY 241 AAATTATCAGAAGATGAAGAAATGAAGTACAGAGCAGATGAATCTCTTAATTCAGA 300
Db 1435 AAATTATCAGAAGATGAAGAAATGAAGTACAGAGCAGATGAATCTCTTAATTCAGA 1494
QY 301 TGGGATGCTCTCAGGTAGTACATGGAAGAAACAAAGCAATTTACATAGATTTTATG 360

Db 1495 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTTAATG 1554
QY 361 GATCTCCAGAAATC-GAAACTGAAGAGCTTGAATGACTGGCTAAACAAAACAGAGAAGA 419
Db 1555 GATCTCCAGAAATCAGAACTGGAAGAGTTGAATGACTGGCTTAACAAAACAGAGAAGA 1614
QY 420 ACAAGGAAATGGAAGAGAGCCCTTTGGACCTGATCTTTGAAGACCTTAAACCGCAAGTA 479
Db 1615 ACAAGGAAATGGAAGAGAGCCCTTTGGACCTGATCTTTGAAGACCTTAAACCGCAAGTA 1674
QY 480 CAACAAATGAAGTGTCTTCAAGAGATCTAGAACAGAACAAAGTCAGGGTCAATTTCTCTC 539
Db 1675 CAACAAATGAAGTGTCTTCAAGAGATCTAGAACAGAACAAAGTCAGGGTCAATTTCTCTC 1734
QY 540 ACTCACATGCTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db 1735 ACTCACATGCTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1794
QY 600 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGC 659
Db 1795 GAACAACCTTAAGGTATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGC 1854
QY 660 TGGGTCTCTTTTACAAGACATCCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 719
Db 1855 TGGGTCTCTTTTACAAGACATCCTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT 1914
QY 720 TTTAGTGCATGGCTTTTCAAGAAAGAGATGCAAGTGAAGAAAGATTCACAACCTGGCTTT 779
Db 1915 TTTAGTGCATGGCTTTTCAAGAAAGAGATGCAAGTGAAGAAAGATTCACAACCTGGCTTT 1974
QY 780 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTA 839
Db 1975 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCGTTTTTAAAGCGGATCTA 2001
QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAAGATCTTCTTCAACA 899
Db 2002 ----- 2001
QY 900 CTGAAGATTAAGTCAGTGAAGCCAGAGACGGAAGCATGGCTGGATTAACCTTTGCCCGGTGT 959
Db 2002 ----- 2001
QY 960 TGGGATAATTTAGTCCAAAACCTTGAAGAGAGTACAGACAGACCCCTTTGAAGAGACTCCAG 1019
Db 2002 ----- 2001
QY 1020 GAACCTCAAGAGGCCACGAGATGAGTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAG 1079
Db 2002 ----- 2001
QY 1080 GGATCCTGGCAGCCCGTGGCGGATCTCTCATTTGACTCTCTCCAAGATCACCTTCGAGAAA 1139
Db 2002 -----CTCGAGAAA 2010
QY 1140 GTCAAGGCACTTCGAGAGAAAATTGGCCCTCTGAAAGAGAACCTGAGCCACGTCATATGAC 1199
Db 2011 GTCAAGGCACTTCGAGAGAAAATTGGCCCTCTGAAAGAGAACCTGAGCCACGTCATATGAC 2070
QY 1200 CTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTCTG 1259
Db 2071 CTTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATTAACCTCAGCACTCTG 2130
QY 1260 GAAGACCTGAACACCAAGATGGAAGCTTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCGAG 1319
Db 2131 GAAGACCTGAACACCAAGATGGAAGCTTTCTGAGGTGGCCGTGAGGACCGAGTCAGGCGAG 2190
QY 1320 CTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTC 1379
Db 2191 CTGCATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTC 2250
QY 1380 CAGGCTCCCTGGGAGAGAGCGCATCTCGCCAAAACAAAGTGGCCCTACTATATCAACACCGAG 1439
Db 2251 CAGGCTCCCTGGGAGAGAGCGCATCTCGCCAAAACAAAGTGGCCCTACTATATCAACACCGAG 2310

D	b	1863	-----AATAGATGAGACCCCTTGAAGAATCCAG	1890
Q	y	1020	GAACTTCAAGAGGCCACGGATGAGCTGGAGCCTCAAAGTCGGCCAAGCTGAGGTGATCAAG	1079
D	b	1891	GAACTTCAAGAGGCCACGGATGAGCTGGAGCCTCAAAGTCGGCCAAGCTGAGGTGATCAAG	1950
Q	y	1080	GGATCTTGAGACCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACTTCGAGAAA	1139
D	b	1951	GGATCTTGAGACCCCGTGGCGATCTCCTCATTTGACTCTCTCCAAAGATCACTTCGAGAAA	2010
Q	y	1140	GTCAGGCATTTGAGAGAGAAATTCGCCTCTGAAAGAGAACGTGAGCCACGCTCAATGAC	1199
D	b	2011	GTCAGGCATTTGAGAGAGAAATTCGCCTCTGAAAGAGAACGTGAGCCACGCTCAATGAC	2070
Q	y	1200	CTTGCTCGCAGCTTACCATTCTTTGGGCAATTCAGCTCTCACGGTATAACCTTCAGCACTCTG	1259
D	b	2071	CTTGCTCGCAGCTTACCATTCTTTGGGCAATTCAGCTCTCACGGTATAACCTTCAGCACTCTG	2130
Q	y	1260	GAAGACCTGAACACCCAGATGGAAGCTTTCTCGAGTGGCCGTCGAGGACCCGAGTCAGGCAG	1319
D	b	2131	GAAGACCTGAACACCCAGATGGAAGCTTTCTCGAGTGGCCGTCGAGGACCCGAGTCAGGCAG	2190
Q	y	1320	CTGCATGAAGCCACACAGGAGCTTTGGTTCAGCATCTCAGACATTTCTTCCAGCTCTGTC	1379
D	b	2191	CTGCATGAAGCCACACAGGAGCTTTGGTTCAGCATCTCAGACATTTCTTCCAGCTCTGTC	2250
Q	y	1380	CAGGGTCCCTTGGGAGAGAGCCATCTCGCCAAAACAAGTGCCCTACTATATCAACACACGAG	1439
D	b	2251	CAGGGTCCCTTGGGAGAGAGCCATCTCGCCAAAACAAGTGCCCTACTATATCAACACAGAG	2310
Q	y	1440	ACTCAAAACAATTGCTGGGACCATCCCAAATGACAGAGCTCTACAGCTTTTAGCTGAC	1499
D	b	2311	ACTCAAAACAATTGCTGGGACCATCCCAAATGACAGAGCTCTACAGCTTTTAGCTGAC	2370
Q	y	1500	CTGAATAATCTCAGATTTCTCAGCTTATAGACATGCCATGAAATCCGGAAGACTGCAGAG	1559
D	b	2371	CTGAATAATCTCAGATTTCTCAGCTTATAGACATGCCATGAAATCCGGAAGACTGCAGAG	2430
Q	y	1560	GCCTTTGCTTGGATCTCTTGAGCCGTGTCAAGCTGCATGTGATGCCCTGGACCAAGCAAC	1619
D	b	2431	GCCTTTGCTTGGATCTCTTGAGCCGTGTCAAGCTGCATGTGATGCCCTGGACCAAGCAAC	2490
Q	y	1620	CTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTTGTTTGACCACTATT	1679
D	b	2491	CTCAAGCAAAATGACAGCCCATGGATATCCTGCAGATTATTAATTTGTTTGACCACTATT	2550
Q	y	1680	TATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACCGTCCCTCTCTCGCTGGATATG	1739
D	b	2551	TATGACCGCCTGGAGCAAGAGCAACAATTTGGTCAACCGTCCCTCTCTCGCTGGATATG	2610
Q	y	1740	TGTCAGAACTGGCTGCTGAATGTTTTATGATACGGGACGAACAGGGAGATCCCGTGTCTTG	1799
D	b	2611	TGTCAGAACTGGCTGCTGAATGTTTTATGATACGGGACGAACAGGGAGATCCCGTGTCTTG	2670
Q	y	1800	TCCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTACAGATAC	1859
D	b	2671	TCCTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAGCAAGTACAGATAC	2730
Q	y	1860	CTTTTCAAGCAAGTGGCAAGTTCAAACAGATTTTGTGACAGCGCAGGCTGGGCCCTCTCTT	1919
D	b	2731	CTTTTCAAGCAAGTGGCAAGTTCAAACAGATTTTGTGACAGCGCAGGCTGGGCCCTCTCTT	2790
Q	y	1920	CTGCATGATTTCTATCCAAATTTCCAAGACAGTTGGGTGAAGTTGCTCATCTTTGGGGCAGT	1979
D	b	2791	CTGCATGATTTCTATCCNAATTTCCAAGACAGTTGGGTGAAGTTGCTCATCTTTGGGGCAGT	2850
Q	y	1980	AACATTGAGCCAAAGTGTCCGGA	2001
D	b	2851	AACATTGAGCCAAAGTGTCCGGA	2872

RESULT 4

AX538620	AX538620	5339 bp	DNA	linear	PAT 23-NOV-2002
LOCUS	AX538620	5339 bp	DNA	linear	PAT 23-NOV-2002
DEFINITION	Sequence 40 from Patent WO0229056.				
ACCESSION	AX538620				
VERSION	AX538620.1	GI:25271166			
KEYWORDS					
SOURCE	Synthetic construct				
ORGANISM	Synthetic construct				
REFERENCE	artificial sequences.				
AUTHORS	1				
TITLE	Chamberlain,J.S. and Harper,S.Q.				
JOURNAL	Mini-dystrophin nucleic acid and peptide sequences				
	Patent: WO 0229056-A 40 11-APR-2002;				
	THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)				
FEATURES	Location/Qualifiers				
source	1..5339				
	/organism="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32630"				
	/note="Synthetic"				
BASE COUNT	1638 a 1191 c 1187 g 1323 t				
ORIGIN					
Query Match	62.7%; Score 1254; DB 6; Length 5339;				
Best Local Similarity	77.9%; Pred. No. 4.9e-304;				
Matches 1621;	Conservative 0; Mismatches 380; Indels 79; Gaps 6;				
Qy	1	GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA	60		
Db	1199	GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA	1258		
Qy	61	GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGAGATTTCTAAT	120		
Db	1259	GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAAGCACAAGGAGATTTCTAAT	1318		
Qy	121	GATGTGGAAGTGTGTGAAGACACAGTTTCTATCTATCGAGGGGTACATGATGGATTTGACA	180		
Db	1319	GATGTGGAAGTGTGTGAAGACACAGTTTCTATCTATCGAGGGGTACATGATGGATTTGACA	1378		
Qy	181	GCCCATCAGGCGCGGTTCGTATATTTCTACAAATTTGGGAAGTAAAGTGTGGAAACAGCA	240		
Db	1379	GCCCATCAGGCGCGGTTCGTATATTTCTACAAATTTGGGAAGTAAAGTGTGGAAACAGCA	1438		
Qy	241	AAATTATCAGAAGATGAAGAAATCGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGA	300		
Db	1439	AAATTATCAGAAGATGAAGAAATCGAAGTACAAGAGCAGATGAATCTCTAAATTCAAGA	1498		
Qy	301	TGGGAATGCTCAGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGATTTTAATG	360		
Db	1499	TGGGAATGCTCAGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGATTTACTG	1558		
Qy	361	GATCTCCAGAATC-GAAACTGAAAGATTTGAATGACTGGCTAACAAAAACAAGAAAGA	419		
Db	1559	CAACAGTTCCTCCCTGGACCTGGAAAGTTTCTTGCTTGCCTACAGAGCTGAACAACT	1618		
Qy	420	ACAAGGAAATGGAGGAAGGCTCTTTGAGCTGATCTTTGAAGACCTTAAACGCCAAGTA	479		
Db	1619	GCCAAATGCTCTACAGGATGCTACCCCTAAGGAAAGGCTCTTGAAGACTCCAGAGGAGTA	1678		
Qy	480	CAACAACTAAGTGCTTCAAGAGATCTAGAACAGACACAGTCAAGGTCATTTCTCTC	539		
Db	1679	AAGAGCTGATGAAACAATGGCAGACCTTCAAGGTGAAATTTGAAGCTCACACAGATGTT	1738		
Qy	540	ACTCACATGGTGGTGTAG-----TTGATGAATCTAGTGGAGATCACGCAACT	587		
Db	1739	TATCAAACTGGATGAAAACAGCCAAAAATCTCTGAGATCCTCGAAGTTCCGATGAT	1798		
Qy	588	GCTGCTTTGGAAGACAACTTAAG---TATTGGGAGATCGATGGCAACATCTGTAGA	644		
Db	1799	GCACTCCTGTTCAAGAGAGCGTTTGGATACATGAATCTCAAGTGGAGTGAATCTCGGAA	1858		
Qy	645	TGACAGAAAGACCGCTGGGTCTTTTTCACAGACATCTCTCTCAAAATGGCAAGCTTACT	704		

Db 1859 AAGTCTCTCAACATTAGTCCATTTGGAGCCAGTCTTGACCACTGGAGCGTCTGCAC 1918
QY 705 GAAGAACAGTGCCTTTTATGTCATGGCTTTTCAGAAAAAAGAGATGCGATGAACAGATT 764
Db 1919 CTTTCTCTCGAGGAACCTTCTGTGTGGCTACAGCTGAAGATGATGAATTAAGCGCGAG 1978
QY 765 CACACAACCTGGCTTTAAAGATCAAAATGAATGTTATCAAGTCTTCAAAAACCTGGCCCTT 824
Db 1979 GCACCTATTGGAGGCGACTTTCCAGCAGTTTCAGAGCAGAACGATGTACATAGGCGCTTC 2038
QY 825 TTAAGAAGCGATCTAGAAAAGAAAAAGCAATCCATGGGCAAACTGTA----- 871
Db 2039 AAGAGGGAATTTGAAAAATTAAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTACGAATA 2098
QY 872 -TTCACTCAAAACAAGATCTTCTTTCAACACTGAAGATAAGTCACTGACCCAGAGACGG 930
Db 2099 TTTCTACAGACGCGCTTTGGNAGACTAGAGAACTCTACAGAGAGCCACAGAGCTG 2158
QY 931 AAGCATGGCTGGA-----TAACTTTGCCCGGTTTGGG 963
Db 2159 CTTCTTGAGGAGAGAGCCAGCAATGTCACTCGCTTTACGAAAGCAGCGCTGAGGAGTTC 2218
QY 964 ATAAATTTAGTCCAAACTTGA-----AAGAGTACACACAG 1001
Db 2219 AATPCTGAGTGGGAAAAATTGAACCTGCACTCCGCTGACGTCGACAGAGAAAAATAGATGAG 2278
QY 1002 ACCCTTTGAAGACTCCAGGAACCTTCAAGAGGCGCCAGGATGAGCTGAGCTCAAGCTGCGC 1061
Db 2279 ACCCTTTGAAGACTCCAGGAACCTTCAAGAGGCGCCAGGATGAGCTGAGCTCAAGCTGCGC 2338
QY 1062 CAAGCTGAGTGTATCAAGGATCTCTGAGCGCGTGGGGATCTCTCATTTGACTCTCTC 1121
Db 2339 CAAGCTGAGTGTATCAAGGATCTCTGAGCGCGTGGGGATCTCTCATTTGACTCTCTC 2398
QY 1122 CAAGATCACTCCAGAAAGTCAAGGACCTTTCAGAGAGAAATTCGCGCTCTGAAAAGAGAAC 1181
Db 2399 CAAGATCACTCCAGAAAGTCAAGGACCTTTCAGAGAGAAATTCGCGCTCTGAAAAGAGAAC 2458
QY 1182 GTGAGCGAGTCAATGACCTTGCTGCCAGCTTACACAGATGGAAGCTTCTGAGGTGGCCCTC 1241
Db 2459 GTGAGCGAGTCAATGACCTTGCTGCCAGCTTACACAGATGGAAGCTTCTGAGGTGGCCCTC 2518
QY 1242 TATAACTCAGCACTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTGGCCCTC 1301
Db 2519 TATAACTCAGCACTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGAGGTGGCCCTC 2578
QY 1302 GAGGACCGAGTCAGGAGCTGCAATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCAC 1361
Db 2579 GAGGACCGAGTCAGGAGCTGCAATGAAGCCCAAGGACTTTGGTCCAGCATCTCAGCAC 2638
QY 1362 TTTCTTTCCAGCTCTGTCAGGCTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGGCC 1421
Db 2639 TTTCTTTCCAGCTCTGTCAGGCTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGGCC 2698
QY 1422 TACTATATCAACACAGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTC 1481
Db 2699 TACTATATCAACACAGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTC 2758
QY 1482 TACGAGCTTTTACGTAACCTGATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAA 1541
Db 2759 TACGAGCTTTTACGTAACCTGATATGTCAGATTTCTCAGCTTATAGGACTGCCATGAA 2818
QY 1542 CTCGGAAGACTGCAGAGGCGCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGCATGTGAT 1601
Db 2819 CTCGGAAGACTGCAGAGGCGCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGCATGTGAT 2878
QY 1602 GCCTTGACACAGACACACTCAGAGAAATGACAGCCCATGATATCTCAGATTTT 1661
Db 2879 GCCTTGACACAGACACACTCAGAGAAATGACAGCCCATGATATCTCAGATTTT 2938
QY 1662 AATTGTTTGACCACTATTTATGACCGCTTGGAGCAAGAGACACAAATTTGTTCAACGTC 1721
Db 2939 AATTGTTTGACCACTATTTATGACCGCTTGGAGCAAGAGACACAAATTTGTTCAACGTC 2998

QY 1722 CCTCTCTGGTGGATATGTCTGTAAGTGGCTGCTGAATGTTTATGATACGGGACGAACA 1781
Db 2999 CCTCTCTGGTGGATATGTCTGTAAGTGGCTGCTGAATGTTTATGATACGGGACGAACA 3058
QY 1782 GGGAGGATCCGTCTGCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTG 1841
Db 3059 GGGAGGATCCGTCTGCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTG 3118
QY 1842 GAAGACAACTGACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCA 1901
Db 3119 GAAGACAACTGACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCA 3178
QY 1902 CGCAGGCTGGGCTCTCTCTGTCATGATTTTATCCAAATTTCCAAAGACAGTTGGGTGA 1961
Db 3179 CGCAGGCTGGGCTCTCTCTGTCATGATTTTATCCAAATTTCCAAAGACAGTTGGGTGA 3238
QY 1962 GCATCCTTTGGGGGCGAGTAACTTGAAGTGGTCCGGA 2001
Db 3239 GCATCCTTTGGGGGCGAGTAACTTGAAGTGGTCCGGA 3278

RESULT 5
AX538621
LOCUS AX538621 5462 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 41 from Patent WO0229056.
ACCESSION AX538621
VERSION AX538621.1 GI:25271168
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chamberlain, J.S. and Harper, S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
Patent: WO 0229056-A 41 11-APR-2002;
JOURNAL THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
source
1..5462
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/notes="Synthetic"

BASE COUNT 1668 a 1225 c 1212 g 1357 t
ORIGIN

Query Match 56.8%; Score 1132.2; DB 6; Length 5462;
Best Local Similarity 73.9%; Pred. No. 2e-273;
Matches 1628; Conservative 0; Mismatches 373; Indels 202; Gaps 7;

QY 1 GGCAGTTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAA 60
Db 1199 GGCAGTTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAGAA 1258
QY 61 GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGAAGCAACAGGAGAGATTTCTAAT 120
Db 1259 GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGAAGCAACAGGAGAGATTTCTAAT 1318
QY 121 GATCTGGAAGTGTGAAAGACCCAGTTTCACTCATGAGGGGTACATGATGATTTGACA 180
Db 1319 GATCTGGAAGTGTGAAAGACCCAGTTTCACTCATGAGGGGTACATGATGATTTGACA 1378
QY 181 GCCCATCAGGCGCGGGTGGTAAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 240
Db 1379 GCCCATCAGGCGCGGGTGGTAAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 1438
QY 241 AAATTATCAGAACTGAGAACTGAGTACAGAGCAGATGAATCTCCTTAATTTCAAGA 300
Db 1439 AAATTATCAGAACTGAGAACTGAGTACAGAGCAGATGAATCTCCTTAATTTCAAGA 1498
QY 301 TGGGAATGCTCAGGGGTAGCTAGCATGGGAAAAACAAAGCAATTTTACATAGAGTTTAAATG 360
Db 1499 TGGGAATGCTCAGGGGTAGCTAGCATGGGAAAAACAAAGCAATTTTACATAGCTCTGACTG 1558

QY 361 ----- 360
Db 1559 ACCACTATTGGAGCCTCTCTACTCAGACTGTACTCTGTGACACAACTGTGTTACT 1618
QY 361 -----GATCTCCAGAACTGAAA----- 377
Db 1619 AAGGAACCTGCCATCTCTCAAACTAGAAATGCCATCTTCCTTGATGTGGAGCATAGATTA 1678
QY 378 -----CTGAAGAGTGTGAATGACTGGCTTAACAAAAACAGAA 416
Db 1679 CTGCAACAGATTCCTCCCTGGACCTTGAAAAGTTTCTTCCTGGCTTACAGAAGCTGAAACA 1738
QY 417 AGAACAAGGAAATGGAGGAGAGCCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAA 476
Db 1739 ACTGCCAATGTCTTACAGAGTGTACCCGTAAGAAAGGCTCTTAGAAGACTCCAAGGGA 1798
QY 477 GTACAAACAATAAGAGTGTCTTCAAGAGATCTAGAAACAAGACAAAGTCAGGGTCAATTTCT 536
Db 1799 GTAAAAGAGCTGATGAACAATGCAAGCTCCAGGTGAATTTGAAGCTCACAGAT 1858
QY 537 CTCACCTACATGGTGGTGTAGTTGATGA-----ATCTAGTGAGATCACGA 584
Db 1859 GTTTATCACAACTGGATGAAAACAGCAAAAATCTTGAGATCCCTGGAAGTTCCGAT 1918
QY 585 ACTGCTGCTTTGGAAGAACAACTTAAGG---TATTGGGAGATCGATGGGCAACATCTGT 641
Db 1919 GATGCACTCTGTTACAAAGACGTTTGATTAACATGAACTTCAAGTGGAGTGAATTCGG 1978
QY 642 AGATGACAGAAGACCGCTGGTCTTTTACAGACATCTCTTCAATGGCAACGCTCTT 701
Db 1979 AAAAAGTCTCTCAACATTAGTCCATTTGGAAGCCAGTTCTGACCAGTGGAGGCTCTG 2038
QY 702 ACTGAAGAACAGTGCCTTTTAGTGATGCTTTTCAGAAAAGAAAGATGAGTGAACAAG 761
Db 2039 CACCTTCTCTGACAGAACTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGCCGG 2098
QY 762 ATTCACACAACCTGCTTTTAAAGATCAAAATGAATGTATTCAAGTCTTCAAAAACCTGGCC 821
Db 2099 CAGCACCTATTGGAGCGACTTTCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGGCC 2158
QY 822 GTTTTAAAGCGGATCTAGAAAAGAAAGCAATCCATGGGCAAACTGTA----- 871
Db 2159 TTCAAGAGGGAATGAAACATAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGA 2218
QY 872 ----TTCACCTCAAAACAGATCTCTTTTCAACACTGAAGATAGTCAGTGACCCGAAGA 927
Db 2219 ATATTCTGACAGACGACCTTTTGGAGGACTAGAGAACTCTACAGGAGCCAGAGAG 2278
QY 928 CGGAAGCATCGCTGGA-----TAACTTTGCCCGGTGT 960
Db 2279 CTGCCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAGCAGGCTGAGGAG 2338
QY 961 GGGATAAATTAGTCCAAAACCTGAA-----AAGATACAGCA 998
Db 2339 GTCAATCTAGTGGGAAAAATTTGAACCTGCATCTCGCTGACTGGCAGAGAAAAATAGAT 2398
QY 999 CAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGACCTCAAGCTG 1058
Db 2399 GAGACCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACCGATGAGCTGACCTCAAGCTG 2458
QY 1059 CGCAGACTGAGTGATCAAGGATCTTGACAGCCGCTGGGCGATCTCTCATTTGACTCT 1118
Db 2459 CGCCAAGCTGAGTGATCAAGGATCTTGACAGCCGCTGGGCGATCTCTCATTTGACTCT 2518
QY 1119 CTCCAAGATCACCTCGAGAAAGTCAAGGACTTTCAGAGGAGAAATTCGCGCTCTGAAAGAG 1178
Db 2519 CTCCAAGATCACCTCGAGAAAGTCAAGGACTTTCAGAGGAGAAATTCGCGCTCTGAAAGAG 2578
QY 1179 AAGTGAGCCAGCTCAATGACTTGTGCGCAGCTTACCACTTTGGGCATTCAGCTCTCA 1238
Db 2579 AAGGTAGCCAGCTCAATGACTTGTGCGCAGCTTACCACTTTGGGCATTCAGCTCTCA 2638

QY 1239 CCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGCC 1298
Db 2639 CCGTATAACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTCTGAGGTGCC 2698
QY 1299 GTGAGGACCGAGTCAAGCAGCTGATGAAGCCACAGGACTTTGGTCCAGCATCTCAG 1358
Db 2699 GTGAGGACCGAGTCAAGCAGCTGATGAAGCCACAGGACTTTGGTCCAGCATCTCAG 2758
QY 1359 CACTTTCTTTCCACCTCTGTCTCAGGCTCCCTGGGAGAGAGCCATCTCCCAAAACAAAGTG 1418
Db 2759 CACTTTCTTTCCACCTCTGTCTCAGGCTCCCTGGGAGAGAGCCATCTCCCAAAACAAAGTG 2818
QY 1419 CCCTACTATATCAACACAGACTCAAAACAATCTGTCTGGGACCATCCCAAAATGACAG 1478
Db 2819 CCCTACTATATCAACACAGACTCAAAACAATCTGTCTGGGACCATCCCAAAATGACAG 2878
QY 1479 CTCTACCACTCTTTAGCTGACTGAATAATCTCAGATTCTCAGCTTATAGACTGCCATG 1538
Db 2879 CTCTACCACTCTTTAGCTGACTGAATAATCTCAGATTCTCAGCTTATAGACTGCCATG 2938
QY 1539 AAATCCGAAAGACTGACAGAGCCCTTTTGTGATCTCTTTAGCCCTGTCAAGTGCATGT 1598
Db 2939 AAATCCGAAAGACTGACAGAGCCCTTTTGTGATCTCTTTAGCCCTGTCAAGTGCATGT 2998
QY 1599 GATGCTTTGGACCAAGCAACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATT 1658
Db 2999 GATGCTTTGGACCAAGCAACCTCAAGCAAAATGACAGCCCATGGATATCTCGAGATT 3058
QY 1659 ATTAATTTTGTGACCACTATTATTATGACCCCTGGAGCAAGAGCAACAATTTGGTCAAC 1718
Db 3059 ATTAATTTTGTGACCACTATTATTATGACCCCTGGAGCAAGAGCAACAATTTGGTCAAC 3118
QY 1719 GTCCCTCTCTCGCTGGATATGTGTCTGAATCTGCTGCTGAATGTATGATACGGGAGA 1778
Db 3119 GTCCCTCTCTCGCTGGATATGTGTCTGAATCTGCTGCTGAATGTATGATACGGGAGA 3178
QY 1779 ACAGGAGGATCCCTGCTCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGACAT 1838
Db 3179 ACAGGAGGATCCCTGCTCTCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGACAT 3238
QY 1839 TTGAAGCAAAAGTACAGATACATCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGTGAC 1898
Db 3239 TTGAAGCAAAAGTACAGATACATCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGTGAC 3298
QY 1899 CAGCGAGGCTGGGCTCTCTCTCATGATTCATCCAAATTCACAGACAGTGGGTGAA 1958
Db 3299 CAGCGAGGCTGGGCTCTCTCTCATGATTCATCCAAATTCACAGACAGTGGGTGAA 3358
QY 1959 GTTCATCTCTTTGGGGCAGTAACATTTGAGCCAAAGTCTCCGGA 2001
Db 3359 GTTCATCTCTTTGGGGCAGTAACATTTGAGCCAAAGTCTCCGGA 3401

RESULT 6
AR304538
LOCUS AR304538 5952 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6544786.
ACCESSION AR304538
VERSION AR304538.1 GI:31693691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5952)
AUTHORS Xiao,X. and Liu,P.X.
TITLE Method and vector for producing and transferring trans-spliced peptides
JOURNAL Patent: US 6544786-A 1 08-APR-2003;
FEATURES Location/Qualifiers
 1..5952
 /organism="unknown"
BASE COUNT 1860 a 1344 c 1410 g 1338 t
ORIGIN

Query Match									
Best Local Similarity 50.1%; Score 1002.6; DB 6; Length 5952;									
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
QY	993	ACAGCAGACCCCTTGAAGAATCCAGGAATCTCAAGAGGCCACGGATGAGCTGACCTC	1052						
DB	3679	ATAGATGAGACCCCTTGAAGAATCCAGGAATCTCAAGAGGCCACGGATGAGCTGACCTC	3738						
QY	1053	AACTGCGCAAGCTGAGTGATCAAGGATCTCTGGCAGCCGCTGGCGATCTCTCTCAT	1112						
DB	3739	AACTGCGCAAGCTGAGTGATCAAGGATCTCTGGCAGCCGCTGGCGATCTCTCTCAT	3798						
QY	1113	GACTCTCTCAAGATCACCTCGAGAAGTCAAGGACATTCGAGGAGAAATTCGCGCTCTG	1172						
DB	3799	GACTCTCTCAAGATCACCTCGAGAAGTCAAGGACATTCGAGGAGAAATTCGCGCTCTG	3858						
QY	1173	AAAGAAAGCTGAGCCACGCTCAATGACCTCTGCCAGCTTACCACCTTGGGATTCAG	1232						
DB	3859	AAAGAAAGCTGAGCCACGCTCAATGACCTCTGCCAGCTTACCACCTTGGGATTCAG	3918						
QY	1233	CTCTCACCGTATAACCTTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCGCGAG	1292						
DB	3919	CTCTCACCGTATAACCTTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCGCGAG	3978						
QY	1293	GTGGCCGTTCAGAGACCGAGTCAAGGACCTGATGAGCCACAGGACCTTTGGTCCAGCA	1352						
DB	3979	GTGGCCGTTCAGAGACCGAGTCAAGGACCTGATGAGCCACAGGACCTTTGGTCCAGCA	4038						
QY	1353	TCTCAGCACTTTCTTCCACGCTGTGCCAGGCTCCTGGGAGAGGACCATTCGCGCAAC	1412						
DB	4039	TCTCAGCACTTTCTTCCACGCTGTGCCAGGCTCCTGGGAGAGGACCATTCGCGCAAC	4098						
QY	1413	AAAGTGCCTTACTATATCAACCAAGGACCTCAAGGACCTTCAAGGAGGACCTTCGCGCAAC	1472						
DB	4099	AAAGTGCCTTACTATATCAACCAAGGACCTCAAGGACCTTCAAGGAGGACCTTCGCGCAAC	4158						
QY	1473	ACAGAGCTTACAGCTTTTAGCTGACCTGATATATGTCAGATTTCTCAGCTTTATAGGACT	1532						
DB	4159	ACAGAGCTTACAGCTTTTAGCTGACCTGATATATGTCAGATTTCTCAGCTTTATAGGACT	4218						
QY	1533	GCCATGAACTTCGGAAGCTGCGAAGGCGCTTTGCTTGGATCTCTTGAGCTGTCAGCT	1592						
DB	4219	GCCATGAACTTCGGAAGCTGCGAAGGCGCTTTGCTTGGATCTCTTGAGCTGTCAGCT	4278						
QY	1593	GCAATGATGCTTGGACAGCACAACCTCAAGCAAAATGACAGCCCATGATATCTCTG	1652						
DB	4279	GCAATGATGCTTGGACAGCACAACCTCAAGCAAAATGACAGCCCATGATATCTCTG	4338						
QY	1653	CAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGAGCACAACAAATTTG	1712						
DB	4339	CAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGAGCACAACAAATTTG	4398						
QY	1713	GTCAAGCTTCCTCTGCGTGGATATGTGTCTGAATCTGGCTGCTGAATTTATGATAGG	1772						
DB	4399	GTCAAGCTTCCTCTGCGTGGATATGTGTCTGAATCTGGCTGCTGAATTTATGATAGG	4458						
QY	1773	GGACGACAGGAGGATCCGTCCTCTTTTAAACTGGCATATTTCCCTGTGTAA	1832						
DB	4459	GGACGACAGGAGGATCCGTCCTCTTTTAAACTGGCATATTTCCCTGTGTAA	4518						
QY	1833	GCAATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTT	1892						
DB	4519	GCAATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTT	4578						
QY	1893	TGTGACAGGCGAGGCTGGGCTCTCTTCTGATGTTCTATCCAAATTCGAAGAGAGTTG	1952						
DB	4579	TGTGACAGGCGAGGCTGGGCTCTCTTCTGATGTTCTATCCAAATTCGAAGAGAGTTG	4638						
QY	1953	GGTGAAGTTGCATCTTTTGGGGGAGTAAACATTTGAGCCAAAGTGTCCGGA	2001						
DB	4639	GGTGAAGTTGCATCTTTTGGGGGAGTAAACATTTGAGCCAAAGTGTCCGGA	4687						

RESULT 7									
AX114289									
LOCUS									
DEFINITION									
AX114289									
ACCESSION									
AX114289.1									
VERSION									
AX114289.1									
KEYWORDS									
SOURCE									
ORGANISM									
Homo sapiens (human)									
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
FEATURES									
source									
1. .5952									
/organism="Homo sapiens"									
/mol type="genomic DNA"									
/db xref="taxon:9606"									
misc_feature									
2897..2898									
/note="S4 junction site"									
misc_feature									
3198..3199									
/note="S2 junction site"									
BASE COUNT									
1860 a 1344 c 1410 g 1338 t									
ORIGIN									
Query Match									
Best Local Similarity 50.1%; Score 1002.6; DB 6; Length 5952;									
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
QY	993	ACAGCAGACCCCTTGAAGAATCCAGGAATCTCAAGAGGCCACGGATGAGCTGACCTC	1052						
DB	3679	ATAGATGAGACCCCTTGAAGAATCCAGGAATCTCAAGAGGCCACGGATGAGCTGACCTC	3738						
QY	1053	AACTGCGCAAGCTGAGTGATCAAGGATCTCTGGCAGCCGCTGGCGATCTCTCTCAT	1112						
DB	3739	AACTGCGCAAGCTGAGTGATCAAGGATCTCTGGCAGCCGCTGGCGATCTCTCTCAT	3798						
QY	1113	GACTCTCTCAAGATCACCTTCGAGAAGTCAAGGACATTCGAGGAGAAATTCGCGCTCTG	1172						
DB	3799	GACTCTCTCAAGATCACCTTCGAGAAGTCAAGGACATTCGAGGAGAAATTCGCGCTCTG	3858						
QY	1173	AAAGAAAGCTGAGCCACGCTCAATGACCTCTGCCAGCTTACCACCTTGGGATTCAG	1232						
DB	3859	AAAGAAAGCTGAGCCACGCTCAATGACCTCTGCCAGCTTACCACCTTGGGATTCAG	3918						
QY	1233	CTCTCACCGTATAACCTTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCGCGAG	1292						
DB	3919	CTCTCACCGTATAACCTTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCGCGAG	3978						
QY	1293	GTGGCCGTTCAGAGACCGAGTCAAGGACCTGATGAGCCACAGGACCTTTGGTCCAGCA	1352						
DB	3979	GTGGCCGTTCAGAGACCGAGTCAAGGACCTGATGAGCCACAGGACCTTTGGTCCAGCA	4038						
QY	1353	TCTCAGCACTTTCTTCCACGCTGTGCCAGGCTCCTGGGAGAGGACCATTCGCGCAAC	1412						
DB	4039	TCTCAGCACTTTCTTCCACGCTGTGCCAGGCTCCTGGGAGAGGACCATTCGCGCAAC	1472						
QY	1413	AAAGTGCCTTACTATATCAACCAAGGACCTCAAGGACCTTCAAGGAGGACCTTCGCGCAAC	1472						
DB	4099	AAAGTGCCTTACTATATCAACCAAGGACCTCAAGGACCTTCAAGGAGGACCTTCGCGCAAC	4158						
QY	1473	ACAGAGCTTACAGCTTTTAGCTGACCTGATATATGTCAGATTTCTCAGCTTTATAGGACT	1532						
DB	4159	ACAGAGCTTACAGCTTTTAGCTGACCTGATATATGTCAGATTTCTCAGCTTTATAGGACT	4218						
QY	1533	GCCATGAACTTCGGAAGCTGCGAAGGCGCTTTGCTTGGATCTCTTGAGCTGTCAGCT	1592						
DB	4219	GCCATGAACTTCGGAAGCTGCGAAGGCGCTTTGCTTGGATCTCTTGAGCTGTCAGCT	4278						
QY	1593	GCAATGATGCTTGGACAGCACAACCTCAAGCAAAATGACAGCCCATGATATCTCTG	1652						
DB	4279	GCAATGATGCTTGGACAGCACAACCTCAAGCAAAATGACAGCCCATGATATCTCTG	4338						
QY	1653	CAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGAGCACAACAAATTTG	1712						
DB	4339	CAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGAGCACAACAAATTTG	4398						
QY	1713	GTCAAGCTTCCTCTGCGTGGATATGTGTCTGAATCTGGCTGCTGAATTTATGATAGG	1772						
DB	4399	GTCAAGCTTCCTCTGCGTGGATATGTGTCTGAATCTGGCTGCTGAATTTATGATAGG	4458						
QY	1773	GGACGACAGGAGGATCCGTCCTCTTTTAAACTGGCATATTTCCCTGTGTAA	1832						
DB	4459	GGACGACAGGAGGATCCGTCCTCTTTTAAACTGGCATATTTCCCTGTGTAA	4518						
QY	1833	GCAATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTT	1892						
DB	4519	GCAATTTGGAAGACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTT	4578						
QY	1893	TGTGACAGGCGAGGCTGGGCTCTCTTCTGATGTTCTATCCAAATTCGAAGAGAGTTG	1952						
DB	4579	TGTGACAGGCGAGGCTGGGCTCTCTTCTGATGTTCTATCCAAATTCGAAGAGAGTTG	4638						
QY	1953	GGTGAAGTTGCATCTTTTGGGGGAGTAAACATTTGAGCCAAAGTGTCCGGA	2001						
DB	4639	GGTGAAGTTGCATCTTTTGGGGGAGTAAACATTTGAGCCAAAGTGTCCGGA	4687						

Db 4219 GCCATGAACTCCGAGACTGACAGAGCCCTTTGCTTGGATCTCTTGAGCCCTGTGAGCT 4278
Qy 1593 GCATGTGATGCTTTGGACCCAGCAAACTCAAGCAAAATCACCAGGCCATGGATATCTCTG 1652
Db 4279 GCATGTGATGCTTTGGACCCAGCAAACTCAAGCAAAATCACCAGGCCATGGATATCTCTG 4338
Qy 1653 CAGATTATTAATTTTGGACCACTATTATGACCGCTGAGCAAGAGCACAACATTTG 1712
Db 4339 CAGATTATTAATTTTGGACCACTATTATGACCGCTGAGCAAGAGCACAACATTTG 4398
Qy 1713 GTCACGCTCCCTCTCTCGTGGATATGCTGAACTGGCTGCTGAATGTTTATGATACG 1772
Db 4399 GTCACGCTCCCTCTCTCGTGGATATGCTGAACTGGCTGCTGAATGTTTATGATACG 4458
Qy 1773 GGACGAACAGGAGGATCCGTGCTCTCTTTTAAACTGGATATCTCTCTGCTGATAA 1832
Db 4459 GGACGAACAGGAGGATCCGTGCTCTCTTTTAAACTGGATATCTCTCTGCTGATAA 4518
Qy 1833 GCACATTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTT 1892
Db 4519 GCACATTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTT 4578
Qy 1893 TGTGACCAAGCGAGGCTGGCCCTCTCTGATGATCTTCTGATGATCTTCTGATGATCTT 1952
Db 4579 TGTGACCAAGCGAGGCTGGCCCTCTCTGATGATCTTCTGATGATCTTCTGATGATCTT 4638
Qy 1953 GGTGAAGTTGCATCCCTTTGGGGCAGTAACATTTGAGCCAGTGTCCGGA 2001
Db 4639 GGTGAAGTTGCATCCCTTTGGGGCAGTAACATTTGAGCCAGTGTCCGGA 4687

RESULT 8
AX538622
LOCUS AX538622 8689 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 42 from Patent WO0229056.
ACCESSION AX538622
VERSION AX538622.1 GI:25271171
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 42 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..8689
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"

BASE COUNT 2721 a 1804 c 1861 g 2303 t
ORIGIN

Query Match 50.1%; Score 1002.6; DB 6; Length 8689;
Best Local Similarity 99.6%; Pred.No. 7.9e-241;
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 993 ACAGCAGACCCCTTGAAGACTCCAGAACTTCAAGAGGCCACGGATGAGCTGGACCTC 1052
Db 3725 ATAGATGAGACCCCTTGAAGACTCCAGAACTTCAAGAGGCCACGGATGAGCTGGACCTC 3784
Qy 1053 AAGCTGGCCAAAGCTGAGGTGATCAAGGATCTCGGAGCCCTGGCGATCTCTCAT 1112
Db 3785 AAGCTGGCCAAAGCTGAGGTGATCAAGGATCTCGGAGCCCTGGCGATCTCTCAT 3844
Qy 1113 GACTCTCTCAAGATCACTTCGAGAACTTCAAGCACTTCGAGGAAATTCGCTCTG 1172
Db 3845 GACTCTCTCAAGATCACTTCGAGAACTTCAAGCACTTCGAGGAAATTCGCTCTG 3904
Qy 1173 AAAGAGAACGTGAGCCAGTCAATGACTTGTCTGCGCAGCTTACCACTTTGGGCATTCAG 1232

Db 3905 AAAGAGAACGTGAGCCAGTCAATGACTTGTCTGCGCAGCTTACCACTTTGGGCATTCAG 3964
Qy 1233 CTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTTCAG 1292
Db 3965 CTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAGCTTCTTCAG 4024
Qy 1293 GTGGCCGTGAGGACCGAGTCAAGGAGTCAATGAAGCCCAAGGAGCTTTGGTCCAGCA 1352
Db 4025 GTGGCCGTGAGGACCGAGTCAAGGAGTCAATGAAGCCCAAGGAGCTTTGGTCCAGCA 4084
Qy 1353 TCTCAGCACTTTCTTCCACGCTCTGTCAGGCTCCCTGGGAGAGGACCATCTCGCCAAAC 1412
Db 4085 TCTCAGCACTTTCTTCCACGCTCTGTCAGGCTCCCTGGGAGAGGACCATCTCGCCAAAC 4144
Qy 1413 AAAGTGCCTACTATATCAACACAGAGACTCAAAACAACTTGTCTGGGACCATTCACAAATG 1472
Db 4145 AAAGTGCCTACTATATCAACACAGAGACTCAAAACAACTTGTCTGGGACCATTCACAAATG 4204
Qy 1473 ACAGAGCTCTACCACTCTTTAGCTGACTGAATATGTCAGATTTCTCAGCTTTATAGACT 1532
Db 4205 ACAGAGCTCTACCACTCTTTAGCTGACTGAATATGTCAGATTTCTCAGCTTTATAGACT 4264
Qy 1533 GCCATGAAACTCCGAAGACTGACAGAGCCCTTTTGTGGATCTCTTGTAGCTGTCAAGCT 1592
Db 4265 GCCATGAAACTCCGAAGACTGACAGAGCCCTTTTGTGGATCTCTTGTAGCTGTCAAGCT 4324
Qy 1593 GCATGTGATGCTTTGGACCCAGCAAACTCAAGCAAAATGACCAAGCCCATGGATATCTCTG 1652
Db 4325 GCATGTGATGCTTTGGACCCAGCAAACTCAAGCAAAATGACCAAGCCCATGGATATCTCTG 4384
Qy 1653 CAGATTATTAATTTTGGACCACTATTATGACCGCTTGGAGCAAGACCAACAAATTTG 1712
Db 4385 CAGATTATTAATTTTGGACCACTATTATGACCGCTTGGAGCAAGACCAACAAATTTG 4444
Qy 1713 GTCAAGCTCCCTCTCTCGTGGATATGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1772
Db 4445 GTCAAGCTCCCTCTCTCGTGGATATGCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4504
Qy 1773 GGACGAACAGGAGGATCCGTGCTCTCTTTTAAACTTGGCATCTTCCCTGTGTAAA 1832
Db 4505 GGACGAACAGGAGGATCCGTGCTCTCTTTTAAACTTGGCATCTTCCCTGTGTAAA 4564
Qy 1833 GCACATTTGGAAGACAAGTACAGATACCTTTTAAAGCAAGTGGCAAGTTCACAGGATTT 1892
Db 4565 GCACATTTGGAAGACAAGTACAGATACCTTTTAAAGCAAGTGGCAAGTTCACAGGATTT 4624
Qy 1893 TGTGACCAAGCGAGGCTGGCCCTCTCTGATGATCTTCAAAATTCACAGACAGTTG 1952
Db 4625 TGTGACCAAGCGAGGCTGGCCCTCTCTGATGATCTTCAAAATTCACAGACAGTTG 4684
Qy 1953 GGTGAAGTTGCATCCCTTTGGGGCAGTAACATTTGAGCCAGTGTCCGGA 2001
Db 4685 GGTGAAGTTGCATCCCTTTGGGGCAGTAACATTTGAGCCAGTGTCCGGA 4733

RESULT 9
AX538624
LOCUS AX538624 11443 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 44 from Patent WO0229056.
ACCESSION AX538624
VERSION AX538624.1 GI:25271175
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 44 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
Location/Qualifiers
1..11443
/organism="synthetic construct"

/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic"

BASE COUNT 3707 a 2339 c 2502 g 2895 t
ORIGIN

Query Match 50.1%; Score 1002.6; DB 6; Length 11443;
Best Local Similarity 99.6%; Pred. No. 8e-241;
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 993 ACAGCACAGACCTTGAAGAGCTCCAGGAACCTCAAGAGGCCACGAGTGAAGTGGACCTC 1052
Db 6479 ATAGATGAGACCTTGAAGAGCTCCAGGAACCTCAAGAGGCCACGAGTGAAGTGGACCTC 6538

QY 1053 AAGCTGCGCCAGCTGAGTGTATCAAGGATCCTGGCAGCCCGTGGCGATCTCTCTCATT 1112
Db 6539 AAGCTGCGCCAGCTGAGTGTATCAAGGATCCTGGCAGCCCGTGGCGATCTCTCTCATT 6598

QY 1113 GACTCTCTCCAGATCAGCTCCAGAGTCAAGGACCTTCAGAGGAGAAATTGGCCCTCTG 1172
Db 6599 GACTCTCTCCAGATCAGCTCCAGAGTCAAGGACCTTCAGAGGAGAAATTGGCCCTCTG 6658

QY 1173 AAAGAGAACGTGAGCCACGCTCAATGACCTTGTCTGCCAGCTTACCACCTTGGGCATTTCAG 1232
Db 6659 AAAGAGAACGTGAGCCACGCTCAATGACCTTGTCTGCCAGCTTACCACCTTGGGCATTTCAG 6718

QY 1233 CTCTCAGCTATTAAGCTCAGACTCTGAGAGACCTTGAACACGAGATGAAGCTTCTGCGAG 1292
Db 6719 CTCTCAGCTATTAAGCTCAGACTCTGAGAGACCTTGAACACGAGATGAAGCTTCTGCGAG 6778

QY 1293 GTGGCCGTCCAGGACCGAGTCAAGGAGCTGCTGAAGAGCCACAGGAGCTTTGGTCCAGCA 1352
Db 6779 GTGGCCGTCCAGGACCGAGTCAAGGAGCTGCTGAAGAGCCACAGGAGCTTTGGTCCAGCA 6838

QY 1353 TCTCAGCACTTTCTTTCCAGCTCTGTCAGGCTCCCTGGGAGAGAGCCATTCGCCCAAAAC 1412
Db 6839 TCTCAGCACTTTCTTTCCAGCTCTGTCAGGCTCCCTGGGAGAGAGCCATTCGCCCAAAAC 6898

QY 1413 AAAGTGCCCTACTATATCAACACAGAGCTCAAAACAACTTCTGGGACCATCCCAAAATG 1472
Db 6899 AAAGTGCCCTACTATATCAACACAGAGCTCAAAACAACTTCTGGGACCATCCCAAAATG 6958

QY 1473 ACAGAGCTCTACAGCTTTTAGCTGACCTGTAATATGTGAGATTTCTCAGCTTATAGGACT 1532
Db 6959 ACAGAGCTCTACAGCTTTTAGCTGACCTGTAATATGTGAGATTTCTCAGCTTATAGGACT 7018

QY 1533 GGCATGAACCTCGAGAGCTGAGAGGCCCTTTGCTTGGATCTCTGAGCTCTGTCAGCT 1592
Db 7019 GGCATGAACCTCGAGAGCTGAGAGGCCCTTTGCTTGGATCTCTGAGCTCTGTCAGCT 7078

QY 1593 GCATGTGATGCTTGGACGAGCACAACTCAAGCAAAATGACCAGCCCATGGATATCTCTG 1652
Db 7079 GCATGTGATGCTTGGACGAGCACAACTCAAGCAAAATGACCAGCCCATGGATATCTCTG 7138

QY 1653 CAGATTAATAATGTTGACCACTATTTATGACCGCTTGGAGCAAGAGCACAAAATTTG 1712
Db 7139 CAGATTAATAATGTTGACCACTATTTATGACCGCTTGGAGCAAGAGCACAAAATTTG 7198

QY 1713 GTCAAGCTCCCTCTCTGGTGATATGTCTGAACTGCTGCTGAATGTTTATGATACG 1772
Db 7199 GTCAAGCTCCCTCTCTGGTGATATGTCTGAACTGCTGCTGAATGTTTATGATACG 7258

QY 1773 GGACGAACAGGAGGATCCGCTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGATAA 1832
Db 7259 GGACGAACAGGAGGATCCGCTCTCTGCTTTTAAACTGGCATCATTTCCCTGTGATAA 7318

QY 1833 GCACATTTGGAGACAGTACAGTACCTTTTCAGAGCAAGTGGCAAGTTCACAGGATTT 1892
Db 7319 GCACATTTGGAGACAGTACAGTACCTTTTCAGAGCAAGTGGCAAGTTCACAGGATTT 7378

QY 1893 TGTGACAGCGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCGAAGACAGTTG 1952
Db 7379 TGTGACAGCGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCGAAGACAGTTG 7438
```

QY 1953 GGTCAAGTTGCTATCCTTTTGGGGCGAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db 7439 GGTCAAGTTGCTATCCTTTTGGGGCGAGTAACATTGAGCCAAAGTGTCCGGA 7487

RESULT 10
AX538627

LOCUS AX538627 12057 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 47 from Patent WO0229056.

ACCESSION AX538627

VERSION AX538627.1 GI:25271181

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Chamberlain, J. S. and Harper, S. O.

TITLE Mini-dystrophin nucleic acid and peptide sequences

JOURNAL Patent: WO 0229055-A 47 11-APR-2002;

THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)

FEATURES Location/Qualifiers

source 1..12057

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="Synthetic"

BASE COUNT 4020 a 2456 c 2790 g 2791 t

ORIGIN

Query Match 50.1%; Score 1002.6; DB 6; Length 12057;

Best Local Similarity 99.6%; Pred. No. 8e-241;

Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 993 ACAGCACAGACCTTGAAGAGCTCCAGGAACCTTCAAGAGGCCACGAGTGAAGTGGACCTC 1052

Db 8993 ATAGATGAGACCTTGAAGAGCTCCAGGAACCTTCAAGAGGCCACGAGTGAAGTGGACCTC 9052

QY 1053 AAGCTGCGCCAGCTGAGTGTATCAAGGATCCTGGCAGCCCGTGGCGATCTCTCTCATT 1112

Db 9053 AAGCTGCGCCAGCTGAGTGTATCAAGGATCCTGGCAGCCCGTGGCGATCTCTCTCATT 9112

QY 1113 GACTCTCTCAAGATCACCTCGAGAAAGTCAAGGACCTTCGAGGAGAAATTGGCGCTCTG 1172

Db 9113 GACTCTCTCAAGATCACCTCGAGAAAGTCAAGGACCTTCGAGGAGAAATTGGCGCTCTG 9172

QY 1173 AAAGAGAACGTGAGCCACGCTCAATGACCTTCTCGCAGCTTACCACCTTGGGCATTTCAG 1232

Db 9173 AAAGAGAACGTGAGCCACGCTCAATGACCTTCTCGCAGCTTACCACCTTGGGCATTTCAG 9232

QY 1233 CTCTCACCGTATAAAGCTCAGCACTCTGGAAGACCTTGAACACCCAGATGGAAGCTTCTGCAG 1292

Db 9233 CTCTCACCGTATAAAGCTCAGCACTCTGGAAGACCTTGAACACCCAGATGGAAGCTTCTGCAG 9292

QY 1293 GTGGCCGTCCAGAGACCGAGTCAAGGAGCTGATGAAGCCCAAGGAGCTTTGGTCCAGCA 1352

Db 9293 GTGGCCGTCCAGAGACCGAGTCAAGGAGCTGATGAAGCCCAAGGAGCTTTGGTCCAGCA 9352

QY 1353 TCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC 1412

Db 9353 TCTCAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC 9412

QY 1413 AAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTCTGGGAGCAATCCCAAAATG 1472

Db 9413 AAAGTGCCCTACTATATCAACCCAGAGACTCAAAACAACTTGTCTGGGAGCAATCCCAAAATG 9472

QY 1473 ACAGAGCTCTTACAGTCTTTTAGCTGACCTGAATATGTTCAGATTTCTCAGCTTATAGGACT 1532

Db 9473 ACAGAGCTCTTACAGTCTTTTAGCTGACCTGAATATGTTCAGATTTCTCAGCTTATAGGACT 9532

QY 1533 GCCATGAAACTCCGAGAGCTGCAGAGGCCCTTTTCTGGATCTCTTTCAGGCTGTGAGCT 1592

Db 9533 GCCATGAAACTCCGAGAGCTGCAGAGGCCCTTTTCTGGATCTCTTTCAGGCTGTGAGCT 9592

QY	1593	GCATGTGATGCTTGGACCAACAACTCAAGCAAAATGACCAAGCCCATGATATCCTG	1652
Db	9593	GCATGTGATGCTTGGACCAACAACTCAAGCAAAATGACCAAGCCCATGATATCCTG	9652
QY	1653	CAGATTTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGACCAACAATTTG	1712
Db	9653	CAGATTTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGACCAACAATTTG	9712
QY	1713	GTCACAGTCCCTCTCTGCGTGGATGATGCTGAACTGGCTGCTGAATGTTATGATACG	1772
Db	9713	GTCACAGTCCCTCTCTGCGTGGATGATGCTGAACTGGCTGCTGAATGTTATGATACG	9772
QY	1773	GGAAGCAAGGAGGATCCGTGCTGCTGCTTTTAAACTGCGATCATTTCCCTGTGTAAA	1832
Db	9773	GGAAGCAAGGAGGATCCGTGCTGCTGCTTTTAAACTGCGATCATTTCCCTGTGTAAA	9832
QY	1833	GCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTT	1892
Db	9833	GCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTT	9892
QY	1893	TGTGACGAGCGAGCTGGGCTCCTTCTGATGATTTATTCACAAATTCACACAGGATTTG	1952
Db	9893	TGTGACGAGCGAGCTGGGCTCCTTCTGATGATTTATTCACAAATTCACACAGGATTTG	9952
QY	1953	GGTGAAGTTCATCTTTGGGGGAGTAAACATTCAGCAAGTGTCCCGA	2001
Db	9953	GGTGAAGTTCATCTTTGGGGGAGTAAACATTCAGCAAGTGTCCCGA	10001
RESULT 11			
LOCUS	HSDMDR	12446 bp	mRNA linear PRI 12-SEP-1993
DEFINITION	Human mRNA for dystrophin.		
ACCESSION	X14298		
VERSION	X14298.1	GI:30845	
KEYWORDS	Dmd gene; Duchenne muscular dystrophy; dystrophin.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Rosenthal, A., Speer, A., Billwitz, H., Cross, G.S., Forrest, S.M. and Davies, K.E.		
TITLE	Two human cDNA molecules coding for the Duchenne muscular dystrophy (DMD) locus are highly homologous		
JOURNAL	Nucleic Acids Res. 17 (13), 5391 (1989)		
MEDLINE	89345106		
PUBMED	2668885		
REFERENCE	2 (bases 1 to 12446)		
AUTHORS	Rosenthal, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-FEB-1989) Rosenthal, A., Akademie der Wissenschaften der Ddr., Zentralinstitut fuer Molekularbiologie, Robert-Rössle Str. 10, 1115 Berlin Buch, DDR		
COMMENT	see also M18533 and M20250 for Dmd seqs.; discrepancies compared to M18533 cDNA were located at x14298 pos. 496, 1772, 1965, 2449, 3687, 4229, 4504, 5075, 5332, 5630 and 7194.		
FEATURES	Location/Qualifiers		
source	1..12446		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="X chromosomal, Xp21."		
	/tissue_type="muscle"		
	/dev_stage="fetal and adult."		
CDS	99..11156		
	/note="dystrophin (AA 1 - 3685)"		
	/codon_start=1		
	/protein_id="CAA32479.1"		
	/db_xref="GI:30846"		
	/db_xref="SWISS-PROT:P11532"		
	/translation="MLWEEVEDCYEREDVQKKTFTKWNQAQSKFGKHENLPSDL		

QY	993	ACAGCACAGACCTTTGAAGACTCCAGAACTTCAAGAGCCACGGATGAGCTGACCTC	1052
Db	8893	ATAGATGAGACCTTTGAAGACTCCAGAACTTCAAGAGCCACGGATGAGCTGACCTC	8942
BASE COUNT	4135 a 2524 c 2876 g 2911 t		
ORIGIN			
Query Match	50.1%; Score 1002.6; DB 9; Length 12446;		
Best Local Similarity	99.6%; Pred. No. 8e-241;		
Matches 1005; Conservative	0; Mismatches 4; Indels 0; Gaps 0;		


```

QY 1773 GGACGAACAGGAGGATCCGTGTCCTCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAA 1832
Db 9773 GGACGAACAGGAGGATCCGTGTCCTCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAA 9832
QY 1833 GCACATTTGGAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGATTT 1892
Db 9833 GCACATTTGGAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGATTT 9892
QY 1893 TGTGACCAAGCAGGCTGGGCTCCTTCTGTCATGATTCATCAAAATTCACAGACAGTTG 1952
Db 9893 TGTGACCAAGCAGGCTGGGCTCCTTCTGTCATGATTCATCAAAATTCACAGACAGTTG 9952
QY 1953 GGTGAAGTTGATCCTTTGGGGGAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db 9953 GGTGAAGTTGATCCTTTGGGGGAGTAACATTGAGCCAAAGTGTCCGGA 10001

RESULT 13
AX538581
LOCUS AX538581 13957 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 1 from Patent WO0229056.
ACCESSION AX538581
VERSION AX538581.1 GI:25271086
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chamberlain,J.S. and Harper,S.Q.
TITLE Mini-dystrophin nucleic acid and peptide sequences
JOURNAL Patent: WO 0229056-A 1 11-APR-2002;
THE REGENTS OF THE UNIVERSITY OF MICHIGAN (US)
FEATURES
source
Location/Qualifiers
1. 13957
/or:organism="Homo sapiens"
/mol:types="genomic DNA"
/db xref="taxon:9606"
BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 50.1%; Score 1002.6; DB 6; Length 13957;
Best Local Similarity 99.6%; Pred. No. 8e-241;
Matches 1005; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 993 ACAGCAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGCCACCGATGAGCTGGACCTC 1052
Db 8993 ATAGATCAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGCCACCGATGAGCTGGACCTC 9052
QY 1053 AAGCTGCCCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCGTGGCGATCTCCTCAT 1112
Db 9053 AAGCTGCCCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCGTGGCGATCTCCTCAT 9112
QY 1113 GACTCTCTCAAGATCACTCGAAGAAAGTCAAGGCACTTCGAGAGAAATTCGCGCTCTG 1172
Db 9113 GACTCTCTCAAGATCACTCGAAGAAAGTCAAGGCACTTCGAGAGAAATTCGCGCTCTG 9172
QY 1173 AAGAGAACGTGAGCCAGCTCAATGACCTTCTGCCAGCTTACCACTTTGGGCATTGAG 1232
Db 9173 AAGAGAACGTGAGCCAGCTCAATGACCTTCTGCCAGCTTACCACTTTGGGCATTGAG 9232
QY 1233 CTCTCCCGTATACCTTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCGAG 1292
Db 9233 CTCTCCCGTATACCTTCAGCACTCTGGAAGACCTGAAACACCAAGATGGAAGCTTCTGCGAG 9292
QY 1293 GTGCCCTCGAGGCCAGCTCAGCAGCTGCATGAAGCCACAGGACATTTGGTCAGCA 1352
Db 9293 GTGCCCTCGAGGCCAGCTCAGCAGCTGCATGAAGCCACAGGACATTTGGTCAGCA 9352
QY 1353 TCTCAGCACTTTCTTTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAAC 1412
Db 9353 TCTCAGCACTTTCTTTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAAC 9412

```

```

QY 1413 AAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATG 1472
Db 9413 AAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGCTGGGACCATCCCAAAATG 9472
QY 1473 ACAGAGCTCTACCACTCTTTAGCTGACCTGAATATATGTCAGATTTCAAGCTTATAGGACT 1532
Db 9473 ACAGAGCTCTACCACTCTTTAGCTGACCTGAATATATGTCAGATTTCAAGCTTATAGGACT 9532
QY 1533 GCATGAAGAACTCCGAAGACTCGAGAAGCCCTTTGCTTTGGATCTCTTGAGCTGTGACGT 1592
Db 9533 GCATGAAGAACTCCGAAGACTCGAGAAGCCCTTTGCTTTGGATCTCTTGAGCTGTGACGT 9592
QY 1593 GCATGTGATGCTTGGACCAACCACTCAAGCAAAATGACACGCCCATGATATCCTG 1652
Db 9593 GCATGTGATGCTTGGACCAACCACTCAAGCAAAATGACACGCCCATGATATCCTG 9652
QY 1653 CAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGCAACAATTTG 1712
Db 9653 CAGATTATTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGCAACAATTTG 9712
QY 1713 GTCAACGTCCTCTCTGCGTGATATGTCGTAACCTGCTGCTGAATGTTTATGATACG 1772
Db 9713 GTCAACGTCCTCTCTGCGTGATATGTCGTAACCTGCTGCTGAATGTTTATGATACG 9772
QY 1773 GGACGAACAGGAGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAA 1832
Db 9773 GGACGAACAGGAGGATCCGTGTCCTGCTTTTAAACTGGCATCATTTCCCTGTGTAAA 9832
QY 1833 GCACATTTGGAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTT 1892
Db 9833 GCACATTTGGAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTT 9892
QY 1893 TGTGACCAAGCAGGCTGGGCTCCTCTGTCATGATTCATCAAAATTCACAGACAGTTG 1952
Db 9893 TGTGACCAAGCAGGCTGGGCTCCTCTGTCATGATTCATCAAAATTCACAGACAGTTG 9952
QY 1953 GGTGAAGTTGATCCTTTGGGGGAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db 9953 GGTGAAGTTGATCCTTTGGGGGAGTAACATTGAGCCAAAGTGTCCGGA 10001

RESULT 14
HUMDYS 13957 bp mRNA linear PRI 25-MAY-2000
LOCUS Homo sapiens dystrophin (DMD) mRNA, complete cds.
DEFINITION M18533 M17154 M18026 M20250
ACCESSION M18533.1 GI:181856
VERSION M18533.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1699)
AUTHORS Koenig,M., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and Kunkel,L.M.
TITLE Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals
JOURNAL Cell 50 (3), 509-517 (1987)
MEDLINE 87273512
PUBMED 3607877
REFERENCE 2 (bases 1678 to 3830)
AUTHORS Hoffman,E.P., Monaco,A.P., Feener,C.C. and Kunkel,L.M.
TITLE Conservation of the Duchenne muscular dystrophy gene in mice and humans
JOURNAL Science 238 (4825), 347-350 (1987)
MEDLINE 88018015
PUBMED 3659917
REFERENCE 3 (bases 1 to 13957)
AUTHORS Koenig,M., Monaco,A.P. and Kunkel,L.M.
TITLE The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein

```

JOURNAL
MEDLINE
PUBMED
COMMENT

Cell 53 (2), 219-226 (1988)
88194521
3282674

On May 25, 2000 this sequence version replaced gi:340693.
Draft entry and computer-readable sequence kindly provided by
M.Koenig, 01-APR-1988 The severity of muscular dystrophy is
determined by the size of the deleted DNA segment. Deletions found
in different patients were from positions 302-2200, 473-1168,
1691-1810, and 1169-3011.

FEATURES

source Location/Qualifiers

1. .13957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.3-p21.1"
/tissue_type="muscle"
/dev_stage="fetus"
/note="G00-119-850"
1. .13957
/gene="DMD"
209. .11266
/gene="DMD"
/codon_start=1
/product="dystrophin"
/protein_id="AA53189.1"
/db_xref="GI:181857"

gene

CDS

BASE COUNT 4602 a 2781 c 3122 g 3452 t
ORIGIN
Query Match 50.1%; Score 1002.6; DB 9; Length 13957;
Best Local Similarity 99.6%; Pred. No. 8e-241;
Matches 1005; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 993 ACAGCAGACCCCTTGAAGACTCCAGGACTTCAAGAGGCCACGGATGAGCTGGACCTC 1052
Db 8993 ATAGATGAGACCCCTTGAAGACTCCAGGACTTCAAGAGGCCACGGATGAGCTGGACCTC 9052
Qy 1053 AAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTGCGAGCCCTCGGCGGATCTCCCTCAT 1112
Db 9053 AAGCTGCGCCAAAGCTGAGGTGATCAAGGGATCTGCGAGCCCTCGGCGGATCTCCCTCAT 9112
Qy 1113 GACTCTCTCAAGATCACCTTCGAGAAAGTCAAGGACTCTGAGAGAGAAATGCGCTCTG 1172
Db 9113 GACTCTCTCAAGATCACCTTCGAGAAAGTCAAGGACTCTGAGAGAGAAATGCGCTCTG 9172
Qy 1173 AAAGAGAACTGAGGACCAAGTCAATGACCTTGTCTGCGAGCTTACCACTTTGGGCAATTCAG 1232
Db 9173 AAAGAGAACTGAGGACCAAGTCAATGACCTTGTCTGCGAGCTTACCACTTTGGGCAATTCAG 9232
Qy 1233 CTCTCACCGTATAACCTCAGCCTCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG 1292
Db 9233 CTCTCACCGTATAACCTCAGCCTCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAG 9292
Qy 1293 GTGCGCGCTGAGAGACCGAGTCAAGGAGCTGATGAGGAGCTGATGAGGAGCTGATGAGGAGCTGATGAGGAG 1352
Db 9293 GTGCGCGCTGAGAGACCGAGTCAAGGAGCTGATGAGGAGCTGATGAGGAGCTGATGAGGAGCTGATGAGGAG 9352
Qy 1353 TCTCAGACATTTCTTCTTCCAGCTCTGTCAGGCTTCCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412
Db 9353 TCTCAGACATTTCTTCTTCCAGCTCTGTCAGGCTTCCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9412
Qy 1413 AAAGTGCCCTACTATATCAACAGAGACTCAACAACTTGTCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
Db 9413 AAAGTGCCCTACTATATCAACAGAGACTCAACAACTTGTCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9472
Qy 1473 ACAGAGCTTACAGAGCTTCTTAGCTGAGCTGATGAGTCAATGAGTCTCAGCTTATAGGACT 1532
Db 9473 ACAGAGCTTACAGAGCTTCTTAGCTGAGCTGATGAGTCAATGAGTCTCAGCTTATAGGACT 9532
Qy 1533 GCCATGAATCTCGAAGAGCTGAGAGAGGCGCTTGTGTGATCTTGTGAGCTTGTGAGCTTGTGAGCT 1592
Db 9533 GCCATGAATCTCGAAGAGCTGAGAGAGGCGCTTGTGTGATCTTGTGAGCTTGTGAGCTTGTGAGCT 9592
Qy 1593 GCATGTCATCTTGGACAGAGCAACCTCAACGAG 1652
Db 9593 GCATGTCATCTTGGACAGAGCAACCTCAACGAG 9652
Qy 1653 CAGATTATTAATTTGTTGACCACTATTATTATGACCGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1712
Db 9653 CAGATTATTAATTTGTTGACCACTATTATTATGACCGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9712

DTXKHMVTENINASWRSIHKVREREAALEETHRLLOQPLDLKFLAWLTAETTA
NVLODATERLLEDKGVKELMKWQDLOQRIEATDVTYHNLDSOKSLRSLGSD
DAVLORLLDNWNNFKSBLRKSJLNRSHLEASDOMKBLHLHLSLOELVWLOKDEL
SRQPLGGDFPAVQKQNDVHRAPKRLKTKYKPEVIMSTLETVRILFTEQPLEGLEKLYQ
EPRELPEERANQVRLRLRQAEVNTWEKLNHLSADWQKIDETILERLQELQEAHD
ELDKLRQAEVTKGSPQVGDLLISLOHLEKVKALGEIAPFKENSHVNDLQADL
TTLGILSPVNLSTLEDLNRWKLQVAVEDEVRLQHEARDFGAPGAFHSTVSQGP
WESAIKPNKVPYVNHETITCTWCHPKMTLYQSLADLNNVRFSAVTAKMLRLRLOKA
LCDDLISLGAACDADLHNLKONDQPMDOIICLITTTVDRLQSHNNLVNPLCVD
MCLNLLNVYDTRGTRIRVLAKFTGILSLCAHLEDKRYLFPKQVASTGFCDDRRLL
GLLLHSIQIPQLGVEVASFSGSNIPEPSVSCFQFANKPEIEAALFLDMRLPEOSM
WNLPLVLRVAAATATKQAKNCKECPPIGRYRSKJHFNVDICQSCFSGRVAAGH
KMHPMVETCTPTTSGEDVRDAKLVNKRPTRKRYFAKPRMGYLPVQVLEGNMET
PVTILNFVPSVSPASSPOLSDHDDTHRIEYASRLAEMENSGSYNDSISNEDID
DEHLIHYQCSLNQDPSLQSPRPAQILITSESEGERELERILAEENRNLOAEY
DRUKOOHEHKLSPSPPEMPTSPQSDRAELIAEAKLLRKHRLRARMOLIEDH
NKOLESOLHRLROLLEPOQAEKVNCTTVSSPSTLSQSDSSQPMLLRVVGSQTSDSM
GEEDLLSPQDFTSTGLEEVEQNLNPFPSRGENTFGKPMREDTM

QY	1713	GTCAACGTCCTCTCTCGTGGATATGTGTCGAACCTGGCTGCTGAATGTTTATGATACG	1772			
Db	9713	GTCAACGTCCTCTCTCGTGGATATGTGTCGAACCTGGCTGCTGAATGTTTATGATACG	9772			
QY	1773	GGACGAACAGGGAGGATCCGTGTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA	1832			
Db	9773	GGACGAACAGGGAGGATCCGTGTCTCTTTTAAAACTGGCATCATTTCCCTGTGTAAA	9832			
QY	1833	GCACATTTGGAAGACAGTAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT	1892			
Db	9833	GCACATTTGGAAGACAGTAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT	9892			
QY	1893	TGTGACACGCGCAGGCTGGGCTCTCTTCTGTCATGTTCTATCCAAATTCACAGACAGTTG	1952			
Db	9893	TGTGACACGCGCAGGCTGGGCTCTCTTCTGTCATGTTCTATCCAAATTCACAGACAGTTG	9952			
QY	1953	GGTGAAGTTGCATCCTTTTGGGGGAGTAACATGAGCCAAAGTGTCCGGA	2001			
Db	9953	GGTGAAGTTGCATCCTTTTGGGGGAGTAACATGAGCCAAAGTGTCCGGA	10001			
RESULT 15						
AR220819						
LOCUS	AR220819	13977 bp	DNA linear PAT 26-SEP-2002			
DEFINITION	Sequence 60 from patent US 6426186.					
ACCESSION	AR220819					
VERSION	AR220819.1 GI:23327696					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 13977)					
AUTHORS	Jones, K.A., Volkmut, W. and Walker, M.G.					
TITLE	Bone remodeling genes					
JOURNAL	Patent: US 6426186-A 60 30-JUL-2002;					
FEATURES	Location/Qualifiers					
source	1. 13977					
BASE COUNT	4596 a	2765 c	3120 g 3453 t 43 others			
ORIGIN	/organisms="unknown"					
Query Match						
Best Local Similarity 99.5%; Score 991.6; DB 6; Length 13977;						
Matches 1005; Conservative 0; Mismatches 4; Indels 1; Gaps 1;						
QY	993	ACAGCAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGCCACCGATGAGCTGGACCTC	1052			
Db	8993	ATAGATGAGACCCCTTGAAGACTCCAGGAACCTTCAAGAGCCACCGATGAGCTGGACCTC	9052			
QY	1053	AAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGACGCCCGTGGCGATCTCCTCATTT	1112			
Db	9053	AAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGACGCCCGTGGCGATCTCCTCATTT	9112			
QY	1113	GACTCTCTCAAGATCACCTCGAGAAAGTCAAGGACTTCGAGGAGAAATTCGGCTCTG	1172			
Db	9113	GACTCTCTCAAGATCACCTCGAGAAAGTCAAGGACTTCGAGGAGAAATTCGGCTCTG	9172			
QY	1173	AAAGAGAACCTGAGCCACGTCATATGACCTTGTCTGCGCAGCTTACCACTTTGGGCATT	1232			
Db	9173	AAAGAGAACCTGAGCCACGTCATATGACCTTGTCTGCGCAGCTTACCACTTTGGGCATT	9232			
QY	1233	CTCTCACCGTATAACCTCAGCCTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGCAG	1292			
Db	9233	CTCTCACCGTATAACCTCAGCCTCTGGAAGACCTTGAACACAGATGGAAGCTTCTGCAG	9292			
QY	1293	GTGGCCGTGAGGACCGAGTCAGGACGTCATGAAGCCCAAGGGACTTTGGTCCAGCA	1352			
Db	9293	GTGGCCGTGAGGACCGAGTCAGGACGTCATGAAGCCCAAGGGACTTTGGTCCAGCA	9352			
QY	1353	TCTCAGCAGCTTTCTTTTCAAGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC	1412			
Db	9353	TCTCAGCAGCTTTCTTTTCAAGCTGTCTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAC	9412			

Search completed: February 1, 2004, 23:51:35
Job time : 7142.37 secs

QY	1413	AAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATG	1472
Db	9413	AAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGTGGGACCATCCCAAAATG	9472
QY	1473	ACAGAGCTCTACCCAGTCTTTAGCTGACCTGAATATGTGCAGATTTCTAGCTTATAGGACT	1532
Db	9473	ACAGAGCTCTACCCAGTCTTTAGCTGACCTGAATATGTGCAGATTTCTAGCTTATAGGACT	9532
QY	1533	GCCATGAAACTCCGAAGACTGCAGAAAGCCCTTTGCTTTGGATCTCTTGAGCCTGTCACT	1592
Db	9533	GCCATGAAACTCCGAAGACTGCAGAAAGCCCTTTGCTTTGGATCTCTTGAGCCTGTCACT	9592
QY	1593	GCATGTGATGCTTGGACGAGCACAACCTCAAGCAAAATGACCGCCCATGGATATCCTG	1652
Db	9593	GCATGTGATGCTTGGACGAGCACAACCTCAAGCAAAATGACCGCCCATGGATATCCTG	9652
QY	1653	CAGATTTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGAGCACAACAATTTG	1712
Db	9653	CAGATTTAATTTGTTGACCACTATTTATGACCGCTGGAGCAAGAGCACAACAATTTG	9712
QY	1713	GTCAACGTCCTCTCTCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACG	1772
Db	9713	GTCAACGTCCTCTCTCGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTTATGATACG	9772
QY	1773	GGACGAACAGGGAGGATCCGTGTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAA	1832
Db	9773	GGACGAACAGGGAGGATCCGTGTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAA	9832
QY	1833	GCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT	1892
Db	9833	GCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT	9892
QY	1893	TGTGACACGCGCAGGCT-GGGCCTCCTTCTGTCATGATTTCTATCCAAATTCCAAGACAGTT	1951
Db	9893	TGTGACACGCGCAGGCTGGGGCCTCCTTCTGTCATGATTTCTATCCAAATTCCAAGACAGTT	9952
QY	1952	GGGTGAAGTTGCATCCTTTTGGGGGACAGTAAACATTTGAGCCAAAGTGTCCGGA	2001
Db	9953	GGGTGAAGTTGCATCCTTTTGGGGGACAGTAAACATTTGAGCCAAAGTGTCCGGA	10002

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 10:13:25 ; Search time 475.292 Seconds
(without alignments)
11364.749 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000

Perfect score: 2001
Sequence: 1 99cagtcattgatgagag.....cattgagcaagtgtccgga 2001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2001	100.0	3446	24	Human dystrophin m
2	1990	99.5	4414	24	Adeno-associated v
3	1820.8	91.0	5417	24	DNA encoding mini-
4	1593	79.6	3510	24	Human dystrophin m
5	1593	79.6	4476	24	Adeno-associated v
6	1332	66.6	4402	21	A rod shortened dy
7	1269.2	63.4	4402	21	A rod shortened dy
8	1263	63.1	3858	24	Human dystrophin m

9	1263	63.1	4825	24	AAD37257	Adeno-associated v
10	1263	63.1	4848	24	AAD37263	Adeno-associated v
11	1263	63.1	5060	24	AAD37264	Adeno-associated v
12	1254	62.7	5339	24	ABK81998	DNA encoding mini-
13	1251.2	62.5	3531	24	AAD37238	Human dystrophin m
14	1251.2	62.5	4498	24	AAD37258	Adeno-associated v
15	1245	62.2	4182	24	AAD37230	Human dystrophin m
16	1245	62.2	5149	24	AAD37255	Adeno-associated v
17	1132.2	56.6	5462	24	ABK81999	DNA encoding mini-
18	1112	55.6	3999	24	AAD37234	Human dystrophin m
19	1112	55.6	4966	24	AAD37256	Adeno-associated v
20	1112	55.6	4990	24	AAD37262	Adeno-associated v
21	1002.6	50.1	1821	24	AAD37241	Human dystrophin r
22	1002.6	50.1	2169	24	AAD37232	Human dystrophin r
23	1002.6	50.1	5952	22	AAD06794	Human dystrophin g
24	1002.6	50.1	8689	24	ABK82000	DNA encoding mini-
25	1002.6	50.1	11058	24	AAD37229	Human dystrophin p
26	1002.6	50.1	11241	24	ABK82005	CDNA encoding huma
27	1002.6	50.1	11443	24	ABK82002	DNA encoding mini-
28	1002.6	50.1	12923	10	AAK90338	Sequence of human
29	1002.6	50.1	13957	24	ABT10904	Human breast cance
30	1002.6	50.1	13957	24	ABS69900	Human dystrophin g
31	1002.6	50.1	13957	24	ABN95786	Gene #2284 used to
32	1002.6	50.1	13957	24	ABK81959	CDNA encoding huma
33	1001	50.0	1434	24	AAD37243	Human dystrophin r
34	991.6	49.6	13977	24	ABS70403	Human bone remodel
35	985.8	49.3	1991	24	AAD37231	Human dystrophin N
36	869.8	43.5	13815	24	ABK81960	CDNA encoding mous
37	869.8	43.5	13815	24	ABT199799	Mouse ischaemic co
38	869.8	43.5	19307	17	AAV17558	Shuttle vector PAD
39	868.2	43.4	13815	19	AAV18885	Mus musculus dyster
40	767	38.3	3275	10	AAV97129	Partial sequence o
41	735.4	36.8	4075	21	AAZ48569	A rod shortened dy
42	731.2	36.5	3747	21	AAZ48566	A rod shortened dy
43	677.6	33.9	3163	21	AAZ48571	A rod shortened dy
44	665	33.2	1667	24	AAD37235	Human dystrophin N
45	567.4	28.4	2005	25	ABT33376	NOVX DNA sequence

ALIGNMENTS

RESULT 1

AAD37242
ID AAD37242 standard; DNA; 3446 BP.

XX AAD37242;

DT 21-AUG-2002 (first entry)

XX Human dystrophin minigene delta3447.

DE Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; BMD; BMD;
KW Becker muscular dystrophy; ds.

XX Homo sapiens.

FN WO200183695-A2.

PD 08-NOV-2001.

PF 27-APR-2001; 2001WO-US13677.

XX 28-APR-2000; 2000US-200777P.

XX (XIAO/) XIAO X.

PI Xiao X;

DR WPI; 2002-049342/06.

XX New dystrophin minigene for treating Duchenne or Becker muscular

PT dystrophin comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -

PS Example 1; Page 53-54; 71pp; English.

The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is human dystrophin minigene delc3447 containing nucleotides 1-1992 (N-terminus, hinge H1 and rods R1, R2 and R3), 8749-10227 (rod R34, hinge H4 and CG domain) and 11047-11058 (dystrophin last 3 amino acids).

SQ Sequence 3446 BP; 1074 A; 766 C; 787 G; 819 T; 0 other;

Query Match 100.0%; Score 2001; DB 24; Length 3446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA	60
1000	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA	1059
61	GTATATTGCTGGCTTCCTTCGTCTGAGGACACATTCGCAAGCACAGGAGAGATTTCTAAT	120
1060	GTATATTGCTGGCTTCCTTCGTCTGAGGACACATTCGCAAGCACAGGAGAGATTTCTAAT	1119
121	GATGTGGAAGTGGTGAAGACCGAGTTTCATCTCATGAGCGGTACATGATGCAATTCGACA	180
1120	GATGTGGAAGTGGTGAAGACCGAGTTTCATCTCATGAGCGGTACATGATGCAATTCGACA	1179
181	GCCCATCAGGCGCGGGTGGTATATTTCTACAAATCGGGAAGTAAAGCTGATTTGGACACGGA	240
1180	GCCCATCAGGCGCGGGTGGTATATTTCTACAAATCGGGAAGTAAAGCTGATTTGGACACGGA	1239
241	AAATTTACAGAAGATGAAGAACTGAAGTACAAGACGACATGAATCTCTAAATTCGAAGA	300
1240	AAATTTACAGAAGATGAAGAACTGAAGTACAAGACGACATGAATCTCTAAATTCGAAGA	1299
301	TGGGAATGCCTCAGGTAGCTAGCATGGANAANAACAAGCAATTTACATAGAGTTTAAATG	360
1300	TGGGAATGCCTCAGGTAGCTAGCATGGANAANAACAAGCAATTTACATAGAGTTTAAATG	1359
361	GATCTCCAGAATCGAAACTGAAAGAGTTCGAATGACCTGGCTAAACAAAAACAGAAGAAAGAA	420
1360	GATCTCCAGAATCGAAACTGAAAGAGTTCGAATGACCTGGCTAAACAAAAACAGAAGAAAGAA	1419
421	CAAGAAATAATGGAGGAGAGCCCTCTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTAC	480
1420	CAAGAAATAATGGAGGAGAGCCCTCTGGACCTGATCTTGAAGACCTTAAAGCCCAAGTAC	1479
481	AAACAACAAGAGTGTCTTCAAGAGATCTAGAACAAGACAAGTCAAGGTCAATTTCTCTCA	540
1480	AAACAACAAGAGTGTCTTCAAGAGATCTAGAACAAGACAAGTCAAGGTCAATTTCTCTCA	1539
541	CTCACATGGTGGTGGTATTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAG	600
1540	CTCACATGGTGGTGGTATTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAG	1599
601	AACAACCTTTAAGGTATGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGCT	660
1600	AACAACCTTTAAGGTATGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCCGCT	1659
661	GGGTTCTTTTACAAGACATCTCTTCTCAATTGGCAACGCTTTACTGAAGAACAGTGCCTTT	720
1660	GGGTTCTTTTACAAGACATCTCTTCTCAATTGGCAACGCTTTACTGAAGAACAGTGCCTTT	1719

QY 1801 CTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACC 1860
DB 2800 CTTTAAACTGGCATATTTCCCTGTGTAAAGCACATTTTGGAAAGACAAGTACAGATACC 2859
QY 1861 TTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACACGCGCAGGCTGGGCTCCTTC 1920
DB 2860 TTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTGACACGCGCAGGCTGGGCTCCTTC 2919
QY 1921 TGCATGATTTCTATCCAAATTCACAGAGTTTGGTGAAGTTGCATCCTTTGGGGGAGTA 1980
DB 2920 TGCATGATTTCTATCCAAATTCACAGAGTTTGGTGAAGTTGCATCCTTTGGGGGAGTA 2979
QY 1981 ACATTGAGCCCAAGTGTCCGGA 2001
DB 2980 ACATTGAGCCCAAGTGTCCGGA 3000

RESULT 2
AAD37260
ID AAD37260 standard; DNA; 4414 BP.
XX
XX
AC AAD37260;
XX
XX 21-AUG-2002 (first entry)
XX
XX Adeno-associated virus vector plasmid, AAV-MCK-3447.
XX
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.
XX
XX Chimeric - Homo sapiens.
XX OS Chimeric - Unidentified.
XX
XX W0200183695-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13677.
XX
XX 28-APR-2000; 2000US-200777P.
XX
XX (XIAO//) XIAO X.
XX
XX XIAO X;
XX
XX WPI; 2002-049342/06.
XX
XX New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -
XX
XX Example 1; Page 65-66; 71pp; English.
XX
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid construct containing human dystrophin minigenes, a muscle
XX creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
XX Sequence 4414 BP; 1255 A; 1075 C; 1086 G; 998 T; 0 other;
XX

Query Match 99.5%; Score 1990; DB 24; Length 4414;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGA 60
DB 1757 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTTAGAAGA 1816
QY 61 GTATTATCGTGGCTTCCTTCTCTGAGGACACATTTGAAGCACAGGAGAGATTTCTAAT 120
DB 1817 GTATTATCGTGGCTTCCTTCTCTGAGGACACATTTGAAGCACAGGAGAGATTTCTAAT 1876
QY 121 GATGTGGAAGTGGTGAAGAGACAGTTTCATCTACTCATGAGGGGTACATGATGATTTGACA 180
DB 1877 GATGTGGAAGTGGTGAAGAGACAGTTTCATCTACTCATGAGGGGTACATGATGATTTGACA 1936
QY 181 GCCCATCAGGGCCGGGTTGGTAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 240
DB 1937 GCCCATCAGGGCCGGGTTGGTAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 1996
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTCTAAATTCAGA 300
DB 1997 AAATTATCAGAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTCTAAATTCAGA 2056
QY 301 TGGGAATCCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG 360
DB 2057 TGGGAATCCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG 2116
QY 361 GATCTCCAGAAATC-GAAACTGAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGA 419
DB 2117 GATCTCCAGAAATCAGAAACTGAAGAGTTGAATGACTGGCTAAACAAAACAGAGAAAGA 2176
QY 420 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTCATCTTGAAGACCTTAAACCGCAAGTA 479
DB 2177 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTCATCTTGAAGACCTTAAACCGCAAGTA 2236
QY 480 CAACAACTAAGGTGCTTCAAGAGATCTAGAACAGAAAGTCAAGTCAAGGTCAATTTCTC 539
DB 2237 CAACAACTAAGGTGCTTCAAGAGATCTAGAACAGAAAGTCAAGTCAAGGTCAATTTCTC 2296
QY 540 ACTCATGTGGTGGTGAATTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 599
DB 2297 ACTCATGTGGTGGTGAATTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGGAA 2356
QY 600 GAACAACTTAAGGTATTTGGGAGATCGATGGCAACATCTGTAGATGGACAGAGACCGC 659
DB 2357 GAACAACTTAAGGTATTTGGGAGATCGATGGCAACATCTGTAGATGGACAGAGACCGC 2416
QY 660 TGGTTCCTTTTACAGACATCCTTCTCAATGGCAACGCTCTTACTGAGACAGTGCCTT 719
DB 2417 TGGTTCCTTTTACAGACATCCTTCTCAATGGCAACGCTCTTACTGAGACAGTGCCTT 2476
QY 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGAGTGAACAGAGATTCACAACTGGCTTT 779
DB 2477 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGAGTGAACAGAGATTCACAACTGGCTTT 2536
QY 780 AAAGATCAAAATGAATGTTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAGCGGATCTA 839
DB 2537 AAAGATCAAAATGAATGTTTATCAAGTCTTCAAAAACCTGGCCGTTTTAAAGCGGATCTA 2596
QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGATTTCATCTCAAAACAGATCTTCTTTCAACA 899
DB 2597 GAAAGAAAAAGCAATCCATGGGCAAACTGATTTCATCTCAAAACAGATCTTCTTTCAACA 2656
QY 900 CTGAAGATAAAGTCAGTGAACCCAGAGACGGAAGATGGCTGGATTAATCTTTGCCGGTGT 959
DB 2657 CTGAAGATAAAGTCAGTGAACCCAGAGACGGAAGATGGCTGGATTAATCTTTGCCGGTGT 2716
QY 960 TGGGATATTTAGTCCCAAAATTTGAAGAGTACAGCAGACAGCCCTTGAAGACTCCAG 1019
DB 2717 TGGGATATTTAGTCCCAAAATTTGAAGAGTACAGCAGACAGCCCTTGAAGACTCCAG 2776
QY 1020 GAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAG 1079
DB 2777 GAACTTCAAGAGGCCAGGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGGTGATCAAG 2836

1080 GGATCTCGGCGAGCCGCTGGCGATCTCTCATTTAGTCTCTCAAGATCACTCGAGAAA 1139
Db
2837 GGATCTGCGAGCCGCTGGCGATCTCTCATTTAGTCTCTCAAGATCACTCGAGAAA 2896
Qy 1140 GTCAGGCACTTCGAGGAGAAATGGCGCTCTGAAAGAGAAAGTGAAGCACTCAATGAC 1199
Db 2897 GTCAGGCACTTCGAGGAGAAATGGCGCTCTGAAAGAGAAAGTGAAGCACTCAATGAC 2956
Qy 1200 CTTGCTCGGCGAGCTTACCACTTTGGGCACTTCAGCTCTCAAGCTATCACTCAAGCACTCTG 1259
Db 2957 CTTGCTCGGCGAGCTTACCACTTTGGGCACTTCAGCTCTCAAGCTATCACTCAAGCACTCTG 3016
Qy 1260 GAAGACTGAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACGAGTCAAGGAG 1319
Db 3017 GAAGACTGAACACAGATGGAAGCTTCTGCAAGTGGCGCTCGAGGACGAGTCAAGGAG 3076
Qy 1320 CTGCAATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTCTG 1379
Db 3077 CTGCAATGAAGCCACAGGAGCTTTGGTCCAGATCTCAGCACTTTCTTTCCAGCTCTG 3136
Qy 1380 CAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACAGAG 1439
Db 3137 CAGGCTCCCTGGGAGAGCCATCTCGCCAAACAAAGTGCCTACTATATCAACACAGAG 3196
Qy 1440 ACTCAACAACTTCGCGGAGACCTCCAAATGACAGAGCTCTACAGCTCTTTAGCTGAC 1499
Db 3197 ACTCAACAACTTCGCGGAGACCTCCAAATGACAGAGCTCTACAGCTCTTTAGCTGAC 3256
Qy 1500 CTGAATAATGTCAATCTCAGCTTATAGACTGCCATGAATCTCGAAGCTCGAGAG 1559
Db 3257 CTGAATAATGTCAATCTCAGCTTATAGACTGCCATGAATCTCGAAGCTCGAGAG 3316
Qy 1560 GCCCTTTGCTTGAATCTTTAGGCTGTGAGTGCATGTGATGCTTGAAGCAGCACAAC 1619
Db 3317 GCCCTTTGCTTGAATCTTTAGGCTGTGAGTGCATGTGATGCTTGAAGCAGCACAAC 3376
Qy 1620 CTCAGCAAAATGACAGCCCTGATATCTCGCAGATTAATTTGTTGACCACTATT 1679
Db 3377 CTCAGCAAAATGACAGCCCTGATATCTCGCAGATTAATTTGTTGACCACTATT 3436
Qy 1680 TATGACCGCTGGAGCAAGACACAACAAATTTGGTCAACGCTCCCTCTCTGCGTGATG 1739
Db 3437 TATGACCGCTGGAGCAAGACACAACAAATTTGGTCAACGCTCCCTCTCTGCGTGATG 3496
Qy 1740 TGTCTGAATGGCTGTGATGTTTATGATAGGACGAGAGGAGGATCGTGTCTG 1799
Db 3497 TGTCTGAATGGCTGTGATGTTTATGATAGGACGAGAGGAGGATCGTGTCTG 3556
Qy 1800 TCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACACATTTGGAAGACAAAGTACAGATAC 1859
Db 3557 TCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACACATTTGGAAGACAAAGTACAGATAC 3616
Qy 1860 CTTTTCAGCAAGTGGCAAGTTCAACAGATTTTGTGACAGGAGGAGGCTGGGCTCCTT 1919
Db 3617 CTTTTCAGCAAGTGGCAAGTTCAACAGATTTTGTGACAGGAGGAGGCTGGGCTCCTT 3676
Qy 1920 CTGATGATTTCTATCCAAATTCACAGAGTTGGGTGAAGTTGATCCCTTTGGGGGAGT 1979
Db 3677 CTGATGATTTCTATCCAAATTCACAGAGTTGGGTGAAGTTGATCCCTTTGGGGGAGT 3736
Qy 1980 AACATTGAGCCAAAGTGTCCGGA 2001
Db 3737 AACATTGAGCCAAAGTGTCCGGA 3758

RESULT 3

ABK81997

ID ABK81997 standard; DNA; 5417 BP.

XX

AC

XX

DT

13-AUG-2002 (first entry)

XX

DE DNA encoding mini-dystrophin protein deltaR4-R23.
XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX Homo sapiens.
OS Synthetic.
XX MO200229056-A2.
PN 11-APR-2002.
PD 04-OCT-2001; 2001WO-US311126.
PF 06-OCT-2000; 2000US-238848P.
PR (UNMI) UNIV MICHIGAN.
PA Chamberlain JS, Harper SQ;
XX WPI; 2002-435334/46.
XX A composition for preparing therapeutic drugs, has a mini-dystrophin
PT peptide comprising a specific number of spectrin-like repeat domains,
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX Disclosure; Fig 12; 145pp; English.
XX The invention describes a composition comprising a mini-dystrophin
CC peptide comprising a spectrin-like repeat domain, where the domain
CC comprises n spectrin-like repeats, and contains no more than n
CC spectrin-like repeats, where n is an even number between 4-24, or a
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
CC peptide or the polynucleotide encoding it is useful as a medicament,
CC for preparing a drug for therapeutic application and in the preparation
CC of a composition for treatment of muscle disease, e.g. Duchenne's
CC muscular dystrophy (DMD). This sequence represents a mini-dystrophin
CC sequence of the invention.
XX Sequence 5417 BP; 1700 A; 1192 C; 1182 G; 1343 T; 0 other;

Query Match 91.0%; Score 1820.8; DB 24; Length 5417;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 2; Indels 157; Gaps 2;
Qy 1 GGCAGTTTCATGTGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1199 GGCAGTTTCATGTGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1258
Qy 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCAAGAGAGATTTCTAAT 120
Db 1259 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCAAGAGAGATTTCTAAT 1318
Qy 121 GATGTGAAGTGTGAAAGACCCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 180
Db 1319 GATGTGAAGTGTGAAAGACCCAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 1378
Qy 181 GCCCATCAGGCGCGGTGGTGTATTTCTTACAAATTCGGAAGTAAAGTGTGGAACAGGA 240
Db 1379 GCCCATCAGGCGCGGTGGTGTATTTCTTACAAATTCGGAAGTAAAGTGTGGAACAGGA 1438
Qy 241 AAATTATCAGAAGTGAAGAACTGAAGTACAAAGACAGATGAATCTCTTAAATCAAGA 300
Db 1439 AAATTATCAGAAGTGAAGAACTGAAGTACAAAGACAGATGAATCTCTTAAATCAAGA 1498
Qy 301 TGGGATGCTCAGGCTAGCTAGCGGAAACAAAGCAATTTACATAGATTTTAAATG 360
Db 1499 TGGGATGCTCAGGCTAGCTAGCGGAAACAAAGCAATTTACATAGATTTTAAATG 1558
Qy 361 GATCTCCAGATC-GAATCTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 419
Db 1559 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTAACAAAAACAGAGAAAGA 1618

QY	420	ACAAGAAAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAAGTA	479
DB	1619	ACAAGAAAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGACCTTAAACCGCAAAGTA	1678
QY	480	CAACAAACATAAGGTGCTTTCAGAGAGATCTAGAACAAAGACNAGTCAGGTCATTTCTCTC	539
DB	1679	CAACAAACATAAGGTGCTTTCAGAGAGATCTAGAACAAAGACNAGTCAGGTCATTTCTCTC	1738
QY	540	ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGTGCTTTGGAA	599
DB	1739	ACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGTGCTTTGGAA	1798
QY	600	GAACAACTTAAAGTATTGGGAGATCGATGGCAACACATCTGTAGATGGACAGAACCCGC	659
DB	1799	GAACAACTTAAAGTATTGGGAGATCGATGGCAACACATCTGTAGATGGACAGAACCCGC	1858
QY	660	TGGTTCTTTTACAAAGACATCCTTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT	719
DB	1859	TGGTTCTTTTACAAAGACATCCTTTCTCAAAATGGCAACGCTCTTACTGAAGAACAGTGCCTT	1918
QY	720	TTTAGTCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAAGATTACACAACCTGGCTTT	779
DB	1919	TTTAGTCATGGCTTTTCAGAAAAAGAGATGCAGTGAACAAGATTACACAACCTGGCTTT	1978
QY	780	AAAGATCAAAATCAAAATGTTTATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTA	839
DB	1979	AAAGATCAAAATGAAATGTTTATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTA	2038
QY	840	GAAGAAGAAAGCAATCCATGGGCAAACTGTATTCTACAAACAGATCTTTTCAACA	899
DB	2039	GAAGAAGAAAGCAATCCATGGGCAAACTGTATTCTACAAACAGATCTTTTCAACA	2098
QY	900	CTGAAGAAATAAATCAGTGACCCAGAGACAGGGAAGCATGGCTGATAACTTTGCCCGTCT	959
DB	2099	CTGAAGAAATAAATCAGTGACCCAGAGACAGGGAAGCATGGCTGATAACTTTGCCCGTCT	2158
QY	960	TGGGATAATTTAGTCCAAAAACCTTGAAAAAGAGTACAGCACAGA-----	1002
DB	2159	TGGGATAATTTAGTCCAAAAACCTTGAAAAAGAGTACAGCACAGATTTTACAGGCTGTCAAC	2218
QY	1003	-----	1002
DB	2219	ACCACCTCAGCCATCATTAACACAGACAACATGTAAATGGAAACAGTAACTACGGTGAACCA	2278
QY	1003	-----	1002
DB	2279	AGGAAACAGATCCTGTGTAAGAGCATGCTCAAGAGGAACCTTCCACCACCACTCCCCAAAAG	2338
QY	1003	-----CCCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAG	1043
DB	2339	AAGAGGCAGATTACTGTGGATCTTTGAAAGACTCCAGGAACCTTCAAGAGGCCACGGATGAG	2398
QY	1044	CTGGACCTCAAGCTGCGGCCAAAGCTGAGGTGATCAAGGGAATCCTTGGCAGCCCGCTGGCGCAT	1103
DB	2399	CTGGACCTCAAGCTGCGGCCAAAGCTGAGGTGATCAAGGGAATCCTTGGCAGCCCGCTGGCGCAT	2458
QY	1104	CTCCTCATTTAGCTCTCTCCAGATCACTTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTT	1163
DB	2459	CTCCTCATTTAGCTCTCTCCAGATCACTTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTT	2518
QY	1164	GCGCCTCTGAAAGAGAAACGTGAGCCACGTCATGACCTTTGCTCGCCAGCTTACCACTTTG	1223
DB	2519	GCGCCTCTGAAAGAGAAACGTGAGCCACGTCATGACCTTTGCTCGCCAGCTTACCACTTTG	2578
QY	1224	GGCATTCAGCTCTCACCGGTATAACCTTCAGACATCTTGGAAAGACCTTGAACACAGATGGAG	1283
DB	2579	GGCATTCAGCTCTCACCGGTATAACCTTCAGACATCTTGGAAAGACCTTGAACACAGATGGAG	2638
QY	1284	CTTCTCAGGTGGCCGCTTCAGAGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACATTT	1343
DB	2639	CTTCTCAGGTGGCCGCTTCAGAGACCGAGTCAGGCAGCTGCATGAAGCCCAACAGGACATTT	2698
Y	1344	GGTCCAGCATCTCAGACATTTCTTTCCACGCTGTCTCCAGGGTCCCTGGGAGAGACCAATC	1403

Db	2699	GGTCCAGACTCAGCACTTTCTTTTCACGTCGTGTCACGGTCCCTGGGAGAGAGCCATC	2758
Qy	1404	TCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGCTGGGACCAT	1463
Db	2759	TCGCCAAACAAAGTGCCCTACTATATCAACACGAGACTCAAAACAACCTTGCTGGGACCAT	2818
Qy	1464	CCGAAATATGACAGAGCTCTACAGTCTTTAGCTGACCTGAAATATATGCAATTTCTAGCT	1523
Db	2819	CCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAAATATATGCAATTTCTAGCT	2878
Qy	1524	TATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGCCCTTTTGCTTGGATCTCTCTCAGC	1583
Db	2879	TATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGCCCTTTTGCTTGGATCTCTCTCAGC	2938
Qy	1584	CTGTCAGCTGCATGTGATGCTTGGACGACACAACTCAAGCAAAATGACCGCCCATG	1643
Db	2939	CTGTCAGCTGCATGTGATGCTTGGACGACACAACTCAAGCAAAATGACCGCCCATG	2998
Qy	1644	GATATCTCGCAGATTTAATATGCTTTGACCACTATTATGACCGCTGGAGCAAGAGCAC	1703
Db	2999	GATATCTCGCAGATTTAATATGCTTTGACCACTATTATGACCGCTGGAGCAAGAGCAC	3058
Qy	1704	AACAAATTTGGTCAACGTCCTCTCTCGTGGATATGTCGTAACCTGGCTGGAATGTT	1763
Db	3059	AACAAATTTGGTCAACGTCCTCTCTCGTGGATATGTCGTAACCTGGCTGGAATGTT	3118
Qy	1764	TATGATACGGAGCAACAGGAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCC	1823
Db	3119	TATGATACGGAGCAACAGGAGGATCCGTGCTCTCTTTTAAACTGGCATCATTTCC	3178
Qy	1824	CTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA	1883
Db	3179	CTGTGTAAAGCACATTTGGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA	3238
Qy	1884	ACAGGATTTTGTACACAGCGCAGGCTGGGCCCTCCTTCTGCATGATTTATCCAAATCCA	1943
Db	3239	ACAGGATTTTGTACACAGCGCAGGCTGGGCCCTCCTTCTGCATGATTTATCCAAATCCA	3298
Qy	1944	AGACAGTTGGGTGAAAGTTCATCCTTTTGGGGGCAGTAACATTTGAGCCAGTGTCCGGA	2001
Db	3299	AGACAGTTGGGTGAAAGTTCATCCTTTTGGGGGCAGTAACATTTGAGCCAGTGTCCGGA	3356
RESULT 4			
AAD37240			
ID AAD37240 standard; DNA; 3510 BP.			
XX	AC	AAD37240;	
XX	XX		
DT	XX	21-AUG-2002 (first entry)	
XX	XX	Human dystrophin minigene delta3510.	
XX	XX	Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;	
KW	KW	adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;	
KW	KW	Becker muscular dystrophy; ds.	
XX	XX		
OS	OS	Homo sapiens.	
XX	XX	WO200183695-A2.	
PN	XX	08-NOV-2001.	
XX	XX		
PD	XX	27-APR-2001; 2001WO-US13677.	
XX	XX	28-APR-2000; 2000US-20077P.	
XX	XX	(XIAO/) XIAO X.	
PA	XX		
XX	XX	Xiao X;	
PI	XX	WPI; 2002-049342/06.	
XX	XX		
DR	XX		

XX New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX Example 1; Page 51-52; 71pp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3510 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1 and R2), 8407-10227 (rods R23 and R24, hinge H4 and
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX SQ Sequence 3510 BP; 1073 A; 787 C; 828 G; 822 T; 0 other;
Query Match 79.6%; Score 1593; DB 24; Length 3510;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1811; Conservative 0; Mismatches 190; Indels 64; Gaps 4;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA 60
DB 1000 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA 1059
QY 61 GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGACGACCAAGGAGAGATTCTTAAT 120
DB 1060 GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTCGACGACCAAGGAGAGATTCTTAAT 1119
QY 121 GATGTGAAGTGTGAAGACCACTTCTACTCATGAGGGGTACATGATGATTGACA 180
DB 1120 GATGTGAAGTGTGAAGACCACTTCTACTCATGAGGGGTACATGATGATTGACA 1179
QY 181 GCCCATCAGGCGCGGTGGTGAATATCTTCAATTTGGGAAGTGAAGTGAAGACAGA 240
DB 1180 GCCCATCAGGCGCGGTGGTGAATATCTTCAATTTGGGAAGTGAAGTGAAGACAGA 1239
QY 241 AATATTACAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAATATCAGA 300
DB 1240 AATATTACAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAATATCAGA 1299
QY 301 TGGGAATGCCTCAGGCTAGCTAGCTAGGAAACAAAGCAATTTACATAGAGTTTAAATG 360
DB 1300 TGGGAATGCCTCAGGCTAGCTAGCTAGGAAACAAAGCAATTTACATAGAGTTTAAATG 1359
QY 361 GATCTCCAGATC-GAAACTGAAGAGTTGAATGACTGCTCAACAAAACAGAGAAAGA 419
DB 1360 GATCTCCAGATCAGAAACTGAAGAGTTGAATGACTGCTCAACAAAACAGAGAAAGA 1419
QY 420 ACAAGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 479
DB 1420 ACAAGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 1479
QY 480 CAACACATGAAGTGCTTCAAGAAAGATCTAGAAACAGAAACAAGTCAAGGTCAATCTCTC 539
DB 1480 CAACACATGAAGTGCTTCAAGAAAGATCTAGAAACAGAAACAAGTCAAGGTCAATCTCTC 1539
QY 540 ACTCATGATGTGTGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTCTCTTGGAA 599
DB 1540 ACTCATGATGTGTGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTCTCTTGGAA 1599
QY 600 GAACTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCCG 659
DB 1600 GAACTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCCG 1659
QY 660 TGGGTCTTTTACAGACATCTTCTCAATGGCAAGCTCTTACTGAAGAACAGTGCCTT 719

DB 1660 TGGGTCTTTTACAGACAGTTCTGACCAAGTGAAGGCTCTGACACCTTTCTCTCAGGAA 1719
QY 720 TTTAGTGCATGGCTTTTTCAGAAAAGAGATGCAGTGAACAGATTACACAACTGGCTTT 779
DB 1720 CTTCTGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGACCCTATTGGAGGC 1779
QY 780 AAAGATCAAAATGAATGTTTATCAAGTCTTCAAAAACCTGGCCGCTTTTAAAGCGGATCTA 839
DB 1780 GACTTTCCAGCAGTTTCAGAAAGCAGACGATGTACATAGGGCCCTTCAAGAGGAATTGAAA 1839
QY 840 GAAAGAAAAGCAATCCATGGGCAAACTGTA-----TTCACTCAAAACAAG 885
DB 1840 ACTAAGAGAACTGTAAATCATGAGTACTCTTTGAGACTGTACGAATATTTCTGACAGACAG 1899
QY 886 ATCTTCTTTCAACACTGAAGAATAAGTCAGTGACCCCAAGAGAGCGAAGCATGGCTGA-- 943
DB 1900 CTTTGAAGGACTAGAGAAACTCTACCAAGAGCCCAAGAGCTGCTCTCTGAGGAGAGA 1959
QY 944 -----TAACTTTGCCGGTGTGGGATAATTTAGTCCAAA 978
DB 1960 GCCCAGAAATGTCACTCGGCTTTACGAAAGCAGGCTGAGGAGGTCAATACGTAGTGGGAA 2019
QY 979 AACTTGA-----AAGAGTACAGACAGACCCCTTGAAGACTC 1016
DB 2020 AATTGAACCTGCACTCCGCTGACTGCGCAGAGAAAATAGATGAGACCCCTTGAAGACTC 2079
QY 1017 CAGGAACCTTCAAGAGGCCACGAGTGAAGTCAAGCTGCGCCAAGCTGAGGTGATC 1076
DB 2080 CAGGAACCTTCAAGAGGCCACGAGTGAAGTCAAGCTGCGCCAAGCTGAGGTGATC 2139
QY 1077 AAGGATCTTGGACGCCGCTGGGGATCTCTCTATTGACTCTCTCCAAGATCACTTCGAG 1136
DB 2140 AAGGATCTTGGACGCCGCTGGGGATCTCTCTATTGACTCTCTCCAAGATCACTTCGAG 2199
QY 1137 AAGTCAAGGCACTTTCAGAGAGAAATTCGCCCTCTGAAAAGAGAACTGAGCCACGTCAT 1196
DB 2200 AAGTCAAGGCACTTTCAGAGAGAAATTCGCCCTCTGAAAAGAGAACTGAGCCACGTCAT 2259
QY 1197 GACTTCTCGCAGCTTACCACTTTTGGGATTCAGTCTCACCGTATACCTCAGCACT 1256
DB 2260 GACTTCTCGCAGCTTACCACTTTTGGGATTCAGTCTCACCGTATACCTCAGCACT 2319
QY 1257 CTGGAACACCTGAACACCAAGTGAAGTCTTCAGCTGCGCTGCGAGGACCCAGTCAAG 1316
DB 2320 CTGGAACACCTGAACACCAAGTGAAGTCTTCGAGTGGCGCTCGAGGACCCAGTCAAG 2379
QY 1317 CAGCTGCATGAAGCCCAAGGAGCTTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCT 1376
DB 2380 CAGCTGCATGAAGCCCAAGGAGCTTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCT 2439
QY 1377 GTCCAGGCTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACCAC 1436
DB 2440 GTCCAGGCTCCCTGGGAGAGAGCCATCTGCCAAACAAAGTGCCTACTATATCAACCAC 2499
QY 1437 GAGACTCAAACTTGTCTGGGACCTTCCCAAATGACAGAGCTCTACCAAGTCTTTAGCT 1496
DB 2500 GAGACTCAAACTTGTCTGGGACCTTCCCAAATGACAGAGCTCTACCAAGTCTTTAGCT 2559
QY 1497 GACTGATTAATGTGAGATTCTCAGCTTATAGAGCTGCGCATGAACTCCGAGAGCTGAG 1556
DB 2560 GACTGATTAATGTGAGATTCTCAGCTTATAGAGCTGCGCATGAACTCCGAGAGCTGAG 2619
QY 1557 AAGGCCCTTTGGCTTGAATCTCTTGGAGCCTGTGAGTGCATGTGATGCTTGGACCAAGCAC 1616
DB 2620 AAGGCCCTTTGGCTTGAATCTCTTGGAGCCTGTGAGTGCATGTGATGCTTGGACCAAGCAC 2679
QY 1617 AACCTCAAGCAAAATGACAGCCCATGGATATCTCTCGAGATTAATTAATTTGTTGACACT 1676
DB 2680 AACCTCAAGCAAAATGACAGCCCATGGATATCTCTCGAGATTAATTAATTTGTTGACACT 2739
QY 1677 ATTTATGACCGCTGGAGAGCAAGACCAACAAATTTGTGTCAGCTCCCTCTCTGCTGGGAT 1736

Db 2740 ATTTAAGACCGCTGGAGCAAGACACAACTTTGGTCAACGTCCTCTCTCTCGTGGAT 2799
AAD37259
QY 1737 ATGTCTGTAAGTGGCTGCTGAAATGTTTATGATACGGGACGAAACAGGAGGATCCGTTGC 1796
Db 2800 ATGTCTGTAAGTGGCTGCTGAAATGTTTATGATACGGGACGAAACAGGAGGATCCGTTGC 2859
QY 1797 CTGTCTTTTAAACTGGGATCATTTCCCTGTGTAAAGCACATTTGGGAAGACAAGTACAGA 1856
Db 2860 CTGTCTTTTAAACTGGGATCATTTCCCTGTGTAAAGCACATTTGGGAAGACAAGTACAGA 2919
QY 1857 TACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTGTGACACAGCGCAGCTGGGGCTC 1916
Db 2920 TACCTTTTCAAGCAAGTGGCAAGTTCACAGAGATTTGTGACACAGCGCAGCTGGGGCTC 2979
QY 1917 CTTCTGATGATTTATCCAAATTCACAGAGTGGGTGAAGTTGCATCCTTTTGGGGG 1976
Db 2980 CTTCTGATGATTTATCCAAATTCACAGAGTGGGTGAAGTTGCATCCTTTTGGGGG 3039
QY 1977 AGTAACATTGAGCCCAAGTGTCCGGA 2001
Db 3040 AGTAACATTGAGCCCAAGTGTCCGGA 3064

RESULT 5
ID AAD37259 standard; DNA; 4476 BP.
XX AAD37259;
AC AAD37259;
DT 21-AUG-2002 (first entry)
XX Adeno-associated virus vector plasmid, AAV-MCK-3510.
XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.
XX Chimeric - Homo sapiens.
XX Chimeric - Unidentified.
XX WO200183695-A2.
XX 08-NOV-2001.
XX 27-APR-2001; 2001WO-US13677.
XX 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
XX XIAO X;
XX WPI; 2002-049342/06.
XX New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a
XX dystrophin gene -
XX Example 1; Page 63-65; 71pp; English.
XX The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified
XX N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid construct containing human dystrophin minigenes, a muscle
XX creatine kinase (MCK) promoter and a small polyA signal sequence.

XX SQ Sequence 4476 BP; 1252 A; 1096 C; 1127 G; 1001 T; 0 other;
Query Match 79.6%; Score 1593; DB 24; Length 4476;
Best Local Similarity 87.7%; Fred. No. 0;
Matches 1811; Conservative 0; Mismatches 190; Indels 64; Gaps 4;
QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 60
Db 1756 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAAAGAA 1815
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTAAT 120
Db 1816 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTTGCAAGCAAGGAGAGATTTCTAAT 1875
QY 121 GATGTGGAAGTGGTCAAAAGACAGGTTTCATCTCATGAGGGGTACATGATGGATTGGACA 180
Db 1876 GATGTGGAAGTGGTCAAAAGACAGGTTTCATCTCATGAGGGGTACATGATGGATTGGACA 1935
QY 181 GCCCATCAGGGCCGGTGGTAAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAAACAGA 240
Db 1936 GCCCATCAGGGCCGGTGGTAAATTTCTCAATTTGGGAAGTAAAGCTGATTTGGAAACAGA 1995
QY 241 AAATTATCAGAGATGAAGAACTGAAGTCAAGAGAGATGAATCTCTCTAAATTCAGA 300
Db 1996 AAATTATCAGAGATGAAGAACTGAAGTCAAGAGAGATGAATCTCTCTAAATTCAGA 2055
QY 301 TGGGAATCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 2056 TGGGAATCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTAAATG 2115
QY 361 GATCTCCAGATC-GAACTCAAGAGTTGAATGACTGGCTAAACAAACAGAGAGAGA 419
Db 2116 GATCTCCAGATCAGAACTGAAGAGTTGAATGACTGGCTAAACAAACAGAGAGAGA 2175
QY 420 ACAAGGAAATGGAGAGAGCCTCTTGGACCTGATCTTTGAAGACCTAAACAGCAAGTA 479
Db 2176 ACAAGGAAATGGAGAGAGCCTCTTGGACCTGATCTTTGAAGACCTAAACAGCAAGTA 2235
QY 480 CAACHAATAAGGTGCTTCAAGAGATCTAGAGCAAGAACAGTCAAGGTCAATTTCTCTC 539
Db 2236 CAACHAATAAGGTGCTTCAAGAGATCTAGAGCAAGAACAGTCAAGGTCAATTTCTCTC 2295
QY 540 ACTCAGATGCTGGTGGTATGATGTAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db 2296 ACTCAGATGCTGGTGGTATGATGTAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2355
QY 600 GAAACAATTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGC 659
Db 2356 GAAACAATTAAGGTATTGGGAGATCGATGGCAAAACATCTGTAGATGGACAGAGACCGC 2415
QY 660 TGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAAGCTTCTACTGAAGAACAGTGCCTT 719
Db 2416 TGGGTTCTTTTACAAGACATCCTTCTCAAAATGGCAAGCTTCTGACACCTTCTCTCGAGAA 2475
QY 720 TTTAGTCATGGCTTTTCAGAAAAAGAGATGCAAGTGAACCAAGATTTCAACAACCTGGCTTT 779
Db 2476 CTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGCACCTATTGGAGGC 2535
QY 780 AAGATCAAAATGAATGTTATCAAGTCTTCAAAACCTGGCCGTTTAAAGCGGATCTA 839
Db 2536 GACTTTCCAGAGTTTCAGAGCAAGACGATGTACATAGGGCCCTTCAAGAGGGGAATGAAA 2595
QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTA-----TTCACTCAAAACAG 885
Db 2596 ACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAG 2655
QY 886 ATCTTTTCAACTGAAGATAAGTCAGTGAACCCAGAGACGGAAGCATGGCTGA-- 943
Db 2656 CTTTGGGAAGCAGTAGAGAACTCTACAGAGAGCCAGAGAGCTCCCTCTGAGGAGAGA 2715
QY 944 -----TAACTTTGCCCGTGTGGGATAAATTTAGTCCAAA 978

Db 2716 GCCAGATGTCACTGGCTTCTACGAAAGCAGGCTGAGGAGTCAATCTAGTGGGAA 2775
QY 979 AACTTGAA-----AAGAGTACAGACAGACACCTTTGAAAGACTC 1016
Db 2776 AAATTGAACCTGCATCCGCTGACTGGCAGAGAAATAGATGAGACCTTTGAAGACTC 2835
QY 1017 CAGGAACCTTCAAGAGCCACCGATAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGATC 1076
Db 2836 CAGGAACCTTCAAGAGCCACCGATAGCTGGACCTCAAGCTGGCCAAAGCTGAGGTGATC 2895
QY 1077 AAGGATCTGCGCAGCCGCTGGCCGATCTCCATTTGACTCTCTCAAGATCACCTCGAG 1136
Db 2896 AAGGATCTGCGCAGCCGCTGGCCGATCTCTCAAGATCACCTCGAG 2955
QY 1137 AAAGTCAAGGCACTTGAAGAGAAATTTGGCTCTTGAAGAGAACTGAGCAGCAGTCAAT 1196
Db 2956 AAAGTCAAGGCACTTGAAGAGAAATTTGGCTCTTGAAGAGAACTGAGCAGCAGTCAAT 3015
QY 1197 GACCTTGTCTGCGCAGCTTACCACTTTGGCATTCTAGCTCTACCGTATTAACCTCAGCACT 1256
Db 3016 GACCTTGTCTGCGCAGCTTACCACTTTGGCATTCTAGCTCTACCGTATTAACCTCAGCACT 3075
QY 1257 CTGGAAGACCTGAACACAGATGGAAGCTTTCTGAGGTGGCGTTCGAGGACCGAGTCAAG 1316
Db 3076 CTGGAAGACCTGAACACAGATGGAAGCTTTCTGAGGTGGCGTTCGAGGACCGAGTCAAG 3135
QY 1317 CAGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCT 1376
Db 3136 CAGCTGCATGAAGCCACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCT 3195
QY 1377 GTCCAGGGTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACAC 1436
Db 3196 GTCCAGGGTCCCTGGGAGAGAGCCATCTCGCAAAACAAAGTCCCTACTATATCAACAC 3255
QY 1437 GAGACTCAAAACAACTTGTGGAGACGATCCCAAAATGACAGAGCTCTACAGTCTTTAGCT 1496
Db 3256 GAGACTCAAAACAACTTGTGGAGACGATCCCAAAATGACAGAGCTCTACAGTCTTTAGCT 3315
QY 1497 GACCTCAATAATGCTCAGATCTCAGCTTATAGACTGCCATCAAACTCGAAGACTCGAG 1556
Db 3316 GACCTCAATAATGCTCAGATCTCAGCTTATAGACTGCCATCAAACTCGAAGACTCGAG 3375
QY 1557 AAGGCCCTTTGCTTGGATCTTTGAGCCCTGTCAGCTGCAATGATGCCCTTGACAGCAC 1616
Db 3376 AAGGCCCTTTGCTTGGATCTTTGAGCCCTGTCAGCTGCAATGATGCCCTTGACAGCAC 3435
QY 1617 AACCTCAAGCAAAATGACAGCCATGATATCTCGAGATTATTAATTTGTTGACCACT 1676
Db 3436 AACCTCAAGCAAAATGACAGCCATGATATCTCGAGATTATTAATTTGTTGACCACT 3495
QY 1677 ATTTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCGTGGAT 1736
Db 3496 ATTTATGACCGCTGGAGCAAGACCAACAATTTGGTCAACGTCCTCTCTGCGTGGAT 3555
QY 1737 ATGTGCTGAACCTGGCTGTGAATTTATGATACGGGACGACAGGAGGATTCGTGTC 1796
Db 3556 ATGTGCTGAACCTGGCTGTGAATTTATGATACGGGACGACAGGAGGATTCGTGTC 3615
QY 1797 CTGCTTTTAAACTGGCATCTTTCCCTGTGTAAAGACATTTTGAAGACAAAGTACAGA 1856
Db 3616 CTGCTTTTAAACTGGCATCTTTCCCTGTGTAAAGACATTTTGAAGACAAAGTACAGA 3675
QY 1857 TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGGCAAGCTGGGCTTC 1916
Db 3676 TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACAGGCAAGCTGGGCTTC 3735
QY 1917 CTCTGTCATGATCTATCCAAATTTCAAGACAGTTGGGTGAAGTTCATCTTTTGGGGGC 1976
Db 3736 CTCTGTCATGATCTATCCAAATTTCAAGACAGTTGGGTGAAGTTCATCTTTTGGGGGC 3795
QY 1977 AGTAACATTTGAGCCAAAGTGTCCCGA 2001
Db 3796 AGTAACATTTGAGCCAAAGTGTCCCGA 3820

RESULT 6

AAZ48568
ID AAZ48568 standard; cDNA to mRNA; 4402 BP.
XX
AC AAZ48568;
XX
DT 31-MAR-2000 (first entry)
XX
DE A rod shortened dystrophin (deltaDysAH3) encoding nucleotide sequence.
XX
KW Muscular dystrophy; rod domain; adeno-associated virus; AAV;
XX
KW dystrophin gene; truncated; ds.
XX
OS Homo sapiens.
XX
PN JPI11318467-A.
XX
PD 24-NOV-1999.
XX
PF 08-MAY-1998; 98JP-0142134.
XX
PR 08-MAY-1998; 98JP-0142134.
XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
PA (KOKU-) KOKURIITSU SEISHIN SHINKEI CENT SOCHO.
XX
DR WPI; 2000-100771/09.
XX
DR P-PSDB; AAY59239.
XX
PT A gene encoding a shortened dystrophin - useful for the treatment of muscular dystrophy
XX
PS Claim 7; Page 21-22; 44pp; Japanese.
XX
CC The invention provides a gene for the treatment of muscular dystrophy
CC having at least one rod repeat structure of hinge 1, hinge 4 and rod
CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene
CC and a gene-introducing medium consisting of an adeno-associated virus
CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin
CC genes can be used for the genetic treatment of muscular dystrophy of low
CC immune reaction. The present sequence represents a rod shortened
CC dystrophin encoding sequence.
SQ Sequence 4402 BP; 1339 A; 984 C; 1010 G; 1069 T; 0 other;
Query Match 66.6%; Score 1332; DB 21; Length 4402;
Best Local Similarity 83.8%; Pred. NO. 0;
Matches 1677; Conservative 0; Mismatches 0; Indels 325; Gaps 2;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 60
Db 1195 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAAGAA 1254
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAGGACGAGATTTCTTAAT 120
Db 1255 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGAGGACGAGATTTCTTAAT 1314
QY 121 GATGTGAAGTGTGAAAGACCACTTCTACTCATGAGGGGTACATGATGGATTGACA 180
Db 1315 GATGTGAAGTGTGAAAGACCACTTCTACTCATGAGGGGTACATGATGGATTGACA 1374
QY 181 GCCCATCAGGCGCGGTGGTAAATATTCTTAATTTGGGAAGTAAAGCTGATGGAACAGGA 240
Db 1375 GCCCATCAGGCGCGGTGGTAAATATTCTTAATTTGGGAAGTAAAGCTGATGGAACAGGA 1434
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 300
Db 1435 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 1494
QY 301 TGGGAATGCCTCAGGCTAGCTAGCATGGAAAAACAAGCAATTTACATGAGTTTTAATG 360

XX PS Claim 7; Page 16-17; 44pp; Japanese.

XX CC The invention provides a gene for the treatment of muscular dystrophy

CC having at least one repeat structure of hinge 1, hinge 4 and rod

CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene

CC and a gene-introducing medium consisting of an adeno-associated virus

CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin

CC genes can be used for the genetic treatment of muscular dystrophy of low

CC immune reaction. The present sequence represents a rod shortened

CC dystrophin encoding sequence.

XX SQ Sequence 4402 BP; 1329 A; 1000 C; 1019 G; 1054 T; 0 other;

Query Match 63.4%; Score 1269.2; DB 21; Length 4402;

Best Local Similarity 82.1%; Pred. No. 0;

Matches 1644; Conservative 0; Mismatches 33; Indels 325; Gaps 3;

QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAAGAA 60

DB 1195 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAAGAA 1254

QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTCTTAAT 120

DB 1255 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCAAGGAGAGATTCTTAAT 1314

QY 121 GATCTGGAAGTGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGATTGACA 180

DB 1315 GATCTGGAAGTGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGATTGACA 1374

QY 181 GCCATCTCAGGGCGGTTGGTATATTTCTACATTTGGGAAGTAAAGTATTCGACACGGA 240

DB 1375 GCCATCTCAGGGCGGTTGGTATATTTCTACATTTGGGAAGTAAAGTATTCGACACGGA 1434

QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATCAAGA 300

DB 1435 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATCAAGA 1494

QY 301 TGGGAATCCCTCAGGTAGTACATGAGAAACAAAGCAATTTACATAGATTATATG 360

DB 1495 TGGGAATCCCTCAGGTAGTACATGAGAAACAAAGCAATTTACATAGATTATATG 1554

QY 361 GATCTCCAGAACTC-GAAACTGAAGAGTGAATGATCTGCTAACAAAAACAGAGAAAGA 419

DB 1555 GATCTCCAGAACTCAGAACTGAAGAGTGAATGATCTGCTAACAAAAACAGAGAAAGA 1614

QY 420 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCAAGTA 479

DB 1615 ACAAGGAAATGGAGGAAGAGCCTCTTGGACCTGATCTTGAAGACCTAAACGCAAGTA 1674

QY 480 CAACACATAGGTGCTTCAAGAGATCTAGACACAGACAGTCAAGGCTCAATTTCTTC 539

DB 1675 CAACACATAGGTGCTTCAAGAGATCTAGACACAGACAGTCAAGGCTCAATTTCTTC 1734

QY 540 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 599

DB 1735 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1794

QY 600 GAACAACTTAAGTATGGAGATCGATGGGCAACATCTGTAGATGGACAGAGCCGC 659

DB 1795 GAACAACTTAAGTATGGAGATCGATGGGCAACATCTGTAGATGGACAGAGCCGC 1829

QY 660 TGGGTTCTTTTACAGACATCTCTCTCAATGGCAACGTTTACTGAAGAACAGTGCCTT 719

DB 1830 ----- 1829

QY 720 TTTAGTGCATGGCTTTTTCAGAAAAAGAGATGCAGTGAACAGATTTCACAACTGGCTTT 779

DB 1830 ----- 1829

QY 780 AAGATCAAAATGAATGTTATCAAGTCTTCAAAAACAGTGGCGTTTAAAGCGGATCTA 839

DB 1830 ----- 1829

QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCACTCAAAACAGATCTTCTTTCAACA 899

DB 1830 ----- 1829

QY 900 CTGAAGAAATAGTCACTGAGTCCAGAGAGAGCAATGGCTGGATTAACCTTTGGCCGGTGT 959

DB 1830 ----- 1862

QY 960 TGGGATAATTTAGTCCAAAAAATTTGAAAGAGTACAGACAGACACCTTGAAGACCTCCAG 1019

DB 1863 ----- 1890

QY 1020 GAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAG 1079

DB 1891 GAACTTCAAGAGGCCACCGATGAGCTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAG 1950

QY 1080 GGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGATCTCTCTCAAGATCACTCTCGAGAAA 1139

DB 1951 GGATCTCTGGCAGCCCGTGGCGGATCTCTCATTTGATCTCTCTCAAGATCACTCTCGAGAAA 2010

QY 1140 GTCAAGGCACCTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCATGAC 1199

DB 2011 GTCAAGGCACCTCGAGGAGAAATTTGGCCCTCTGAAAGAGAACTGAGCCACGTCATGAC 2070

QY 1200 CTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTG 1259

DB 2071 CTGCTCGCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATTAACCTCAGCACTCTG 2130

QY 1260 GAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGCTCGAGGACCGAGTCAGGCAG 1319

DB 2131 GAAGACCTGAACACCAAGATGGAAGCTTCTGAGGTGGCCGCTCGAGGACCGAGTCAGGCAG 2190

QY 1320 CTGATGAAGGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTC 1379

DB 2191 CTGATGAAGGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTC 2250

QY 1380 CAGGTCCTCTGGGAGAGAGCCATCTGCGCAAAACAAAGTGCCTTACTATATCAACACAGAG 1439

DB 2251 CAGGTCCTCTGGGAGAGAGCCATCTGCGCAAAACAAAGTGCCTTACTATATCAACACAGAG 2310

QY 1440 ACTCAAAACAATTGCTGGGACCAATCCCAAAATGACAGAGCTCTACAGCTTTTAGTGTAC 1499

DB 2311 ACTCAAAACAATTGCTGGGACCAATCCCAAAATGACAGAGCTCTACAGCTTTTAGTGTAC 2370

QY 1500 CTGAATAATGTCAAGTCTCAGCTTATAGACTGCGCATGAACTCCGAGACTGCGAGAG 1559

DB 2371 CTGAATAATGTCAAGTCTCAGCTTATAGACTGCGCATGAACTCCGAGACTGCGAGAG 2430

QY 1560 GCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGCTGATGCTTGGACCAAGCAAC 1619

DB 2431 GCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGCTGATGCTTGGACCAAGCAAC 2490

QY 1620 CTCAAGCAAAATGACACAGCCCATGATCTCTGAGATTTAATTTGTTGACCACTAT 1679

DB 2491 CTCAAGCAAAATGACACAGCCCATGATCTCTGAGATTTAATTTGTTGACCACTAT 2550

QY 1680 TATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAAGTCTCTGCGTGGATATG 1739

DB 2551 TATGACCGCTGGAGCAAGAGCAACAATTTGGTCAAAGTCTCTGCGTGGATATG 2610

QY 1740 TGTCTGAACCTGGCTGTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGCTG 1799

DB 2611 TGTCTGAACCTGGCTGTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGCTG 2670

QY 1800 TCTTTTAAACTGGCATCTTCCCTGTGTAAGACATTTTGAAGACATTTTGAAGACAGTACATAC 1859

DB 2671 TCTTTTAAACTGGCATCTTCCCTGTGTAAGACATTTTGAAGACATTTTGAAGACAGTACATAC 2730

QY 1860 CTTTTCAGCAAGTGGCAAGTTCAAAGGATTTTGTCAACAGCGCGAGGCTGGCCCTCTT 1919

DB 2731 CTTTTCAGCAAGTGGCAAGTTCAAAGGATTTTGTCAACAGCGCGAGGCTGGCCCTCTT 2790

QY 1920 CTGCATGATCTTATCCAAATCCAAAGACAGTTCGGGTGAAGTTCATCTCTTTGGGGCAGT 1979
AAD37237
Db 2791 CTGCATGATCTTATCCAAATCCAAAGACAGTTCGGGTGAAGTTCATCTCTTTGGGGCAGT 2850
QY 1980 AACATTGAGCCAGTTCGCGA 2001
Db 2851 AACATTGAGCCAGTTCGCGA 2872
RESULT 8
ID AAD37237 standard; DNA; 3858 BP.
XX
AC AAD37237;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo. sapiens.
XX
PN W0200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 48-49; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3849 containing nucleotides 1-1668 (N-terminus,
CC hinge H1 and rods R1, R2), 8059-10227 (rods R22, R23 and R24, hinge H4
CC and CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3858 BP; 1189 A; 866 C; 905 G; 898 T; 0 other;
Query Match 63.1%; Score 1263; DB 24; Length 3858;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;
QY 1 GGCAGTTCATGATGGAGAGTGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1000 GGCAGTTCATGATGGAGAGTGAAGTAACTGGACCGTTATCAACAGCTTTAGAGAA 1059
QY 61 GTATTATCGTGGCTCTTTCTCTGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTCTTTCTCTGCTGAGGACACATTGCAAGCACAAGAGAGATTTCTAAT 1119

QY 121 GATGTGAAGTGGTGAAGACACAGTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 180
Db 1120 GATGTGAAGTGGTGAAGACACAGTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 1179
QY 181 GCCCATCAGGGCCGGGTGGTAAATATTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 240
Db 1180 GCCCATCAGGGCCGGGTGGTAAATATTCTCAATTTGGGAAGTAAAGCTGATTTGGAACAGGA 1239
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAAATTCAGA 300
Db 1240 AAATTATCAGAAGATGAAGAACTGAAGTACAAGACAGATGAATCTCCTAAATTCAGA 1299
QY 301 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAGAAAACAAAGCAATTTACATAGATTTTAATG 360
Db 1300 TGGGAATGCCCTCAGGGTAGCTAGCATGGAAGAAAACAAAGCAATTTACATAGATTTTAATG 1359
QY 361 GATCTCCAGATATC-GAAACTGAAGAGTTGAATCACTGGCTTAACAAAAACAGAGAAGA 419
Db 1360 GATCTCCAGATATCAGAACTGAAGAGTTGAATCACTGGCTTAACAAAAACAGAGAAGA 1419
QY 420 ACAAGGAAATGGAGGAAGAGCCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 479
Db 1420 ACAAGGAAATGGAGGAAGAGCCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA 1479
QY 480 CAACAACATAGGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAGGTCAATCTCTC 539
Db 1480 CAACAACATAGGTGCTTCAAGAGATCTAGAACAGAAACAAGTCAGGTCAATCTCTC 1539
QY 540 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGTCTGCTTTGGAA 599
Db 1540 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGTCTGCTTTGGAA 1599
QY 600 GAACAACATTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGC 659
Db 1600 GAACAACATTAAGGTATTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAGACCGC 1659
QY 660 TGGGTCTCTTTTACAAGACATCCTTCTCAATGGCAACGTCTTACTGAAGAACACGTGCCTT 719
Db 1660 TGGGTCTCTTTTACAAGACATCCTAGATTACTGCAACAGTTCCTCCCTGGACCTGGAAAAG 1719
QY 720 TTTAGTCATCGCTTTTCAAGAAAAGATGCAAGTGAACAAAGATTCACAACTGGCTTT 779
Db 1720 TTTCTTGCTGCTTTACAGAAGCTGAAACCAACTGCCAATGTCTCTACAGGATGCTACCCGT 1779
QY 780 A-----BAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGGCCGTT 824
Db 1780 AAGGAAGCTCTTAGAAGACTCCAGGGAGTAAAGAGCTGATGAACCAATGGCAAGAC 1839
QY 825 TTAAGAGCGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACA 884
Db 1840 CTCCAAGGTGAATTTGAAGCTCACACAGATGTTTATCAACCTGGATGAACACAGCAA 1899
QY 885 GATCTCTTTCAACACTGAAGAATAAGTCAAGTACCAGAGACGGAGGATGGCTGGAT 944
Db 1900 AAAATCTCTGAGATCCCTGGAGGTTCCGATGATGATGATGATGATGATGATGATGAT 1959
QY 945 AACTTTGCCCGTCTGGGATAAATTTAGTCCAAAACCTTGAAGAGAGTACAGCAC----- 999
Db 1960 AACTGAACCTCAAGTGGAGTGAATTCGGAAAAGTCTCTCAACATTAGTCCCATTTG 2019
QY 1000 ----- 999
Db 2020 GAAGCCAGTTCTGACCAGTGGAGCGTCTGCACCTTTCTCTGCGAGAACTTCTGGTGG 2079
QY 1000 ----- 999
Db 2080 CTACAGCTGAAGAAGATGATGAATTAAGCGGCGAGCCACTTATTGGAGCGGACTTTCAGCA 2139
QY 1000 ----- 999
Db 2140 GTTCAGAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACCTAAAGAACCT 2199

```
QY 1000 ----- 999
Db 2200 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGAGCCCTTTGGAAGGA 2259
QY 1000 ----- 999
Db 2260 CTAGAGAAACTCTACAGGAGCCAGAGAGTGCCTCTCTGAGGAGAGCCGAGAATGTC 2319
QY 1000 ----- 999
Db 2320 ACTCGGCTTCTACGAAGCAGCGCTGAGGAGGTCAATCTAGTGGGAAAATTTGAACCTG 2379
QY 1000 -----AGACCTTTGAAGACTCCAGGAACCTTCAA 1028
Db 2380 CACTCGGCTGACTGGCAGAGAAAATAGATGAGACCTTTGAAGACTCCAGGAACCTTCAA 2439
QY 1029 GAGGCCACGAGTGGAGCTGAGCTCAAGCTGCGCAGAGCTGAGTCAATCAAGGATCTCTGG 1088
Db 2440 GAGGCCACGAGTGGAGCTGAGCTCAAGCTGCGCAGAGCTGAGTCAATCAAGGATCTCTGG 2499
QY 1089 CAGCCCGTGGCGAATCTCTCATTTGACTCTCTCCAGAGTCACTCCAGAAAGTCAAGGCA 1148
Db 2500 CAGCCCGTGGCGAATCTCTCATTTGACTCTCTCCAGAGTCACTCCAGAAAGTCAAGGCA 2559
QY 1149 CTTGAGGAGAAATTCGGCTCTGGAAGAGAGCTGAGCCAGTCAATGACTCTCTCTGC 1208
Db 2560 CTTGAGGAGAAATTCGGCTCTGGAAGAGAGCTGAGCCAGTCAATGACTCTCTCTGC 2619
QY 1209 CAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATACTCAGCACTCTGGAAGACCTG 1268
Db 2620 CAGCTTACCACCTTTGGGCAATTCAGCTCTCACCGTATACTCAGCACTCTGGAAGACCTG 2679
QY 1269 AACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGGAGCTGCATGAA 1328
Db 2680 AACACAGATGGAAGCTTCTGAGGTGGCGCTGAGGACCGAGTCAAGGAGCTGCATGAA 2739
QY 1329 GCGCCAGGAGCTTTGGTCTCAGCATCTCAGCACTTCTTTCACGCTCTGTCCAGGCTCC 1388
Db 2740 GCGCCAGGAGCTTTGGTCTCAGCATCTCAGCACTTCTTTCACGCTCTGTCCAGGCTCC 2799
QY 1389 TGGGAGAGGCCATCTGGCCAAACAAAGTGCCTACTATATCAACAGAGACTCAAAACA 1448
Db 2800 TGGGAGAGGCCATCTGGCCAAACAAAGTGCCTACTATATCAACAGAGACTCAAAACA 2859
QY 1449 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTACGCTGACCTGAATAAT 1508
Db 2860 ACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACAGCTTTTACGCTGACCTGAATAAT 2919
QY 1509 GTCAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTCGAAGGCCCTTTGC 1568
Db 2920 GTCAGATTTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTCGAAGGCCCTTTGC 2979
QY 1569 TTGGATCTCTGAGCTGTGAGCTGATGATGCTTGGACAGCAGCAACCTCAAGCAA 1628
Db 2980 TTGGATCTCTGAGCTGTGAGCTGATGATGCTTGGACAGCAGCAACCTCAAGCAA 3039
QY 1629 AATGACAGGCCATGATATCTCTGAGATTAATTAATTTTACCACTATTTATGACCGC 1688
Db 3040 AATGACAGGCCATGATATCTCTGAGATTAATTAATTTTACCACTATTTATGACCGC 3099
QY 1689 CTGGAGCAGAGACAACTAATTTGGTCAACGTCCTCTCTGGGTGATATGTCGTGAAC 1748
Db 3100 CTGGAGCAGAGACAACTAATTTGGTCAACGTCCTCTCTGGGTGATATGTCGTGAAC 3159
QY 1749 TGGCTCTGATGCTTTATGATACGGACGAAACAGGAGGAGTCCGTCTCTGCTTTTAA 1808
Db 3160 TGGCTCTGATGCTTTATGATACGGACGAAACAGGAGGAGTCCGTCTCTGCTTTTAA 3219
QY 1809 ACTGGCATCATTTTCCCTGTGTAAAGCACATTTTGAAGACAACTACAGTACCTTTTCAAG 1868
Db 3220 ACTGGCATCATTTTCCCTGTGTAAAGCACATTTTGAAGACAACTACAGTACCTTTTCAAG 3279
QY 1869 CAAGTGGCAAGTTTCAACAGGATTTTGTGACAGCGCAGGCTGGGCTCTCTCTGATGAT 1928
```

```
Db 3280 CAAGTGCAAGTTCAACAGGATTTGTGACCGAGCTGGGCTCTCTTCTGCAATGAT 3339
QY 1929 TCTATCCAAATTTCCAGACAGTTGGGTGAGTTGCACTCTTTGGGGGCGAGTACATTGAG 1988
Db 3340 TCTATCCAAATTTCCAGACAGTTGGGTGAGTTGCACTCTTTGGGGGCGAGTACATTGAG 3399
QY 1989 CCAAGTGTCGGA 2001
Db 3400 CCAAGTGTCGGA 3412

RESULT 9
AAD37257
ID AAD37257 standard; DNA; 4825 BP.
XX
AC AAD37257;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus vector plasmid, AAV-MCK-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
PN W0200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
PI Xiao X;
XX
WPI; 2002-049342/06.
XX
New dystrophin minigene for treating Duchenne or Becker muscular
dystrophy comprises an N-terminal domain or modified N-terminal domain,
rod repeats, H1 and H4 domains and a cysteine rich domain of a
dystrophin gene -
XX
Example 1; Page 61-62; 71pp; English.
The present invention relates to an isolated nucleotide sequence encoding
a dystrophin minigene. The minigene comprises N-terminal or modified
N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
domains and cysteine-rich domains of dystrophin or utrophin genes. The
invention also relates to a recombinant adeno-associated virus (AAV)
comprising dystrophin minigene operably linked to an expression control
element. The dystrophin minigene in operable linkage with an expression
control element, in a recombinant adeno-associated virus or retrovirus is
useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
dystrophy (BMD) in a mammalian subject. The present sequence is AAV
vector plasmid construct containing human dystrophin minigenes, a muscle
creatine kinase (MCK) promoter and a small polyA signal sequence.
XX
Sequence 4825 BP; 1369 A; 1175 C; 1204 G; 1077 T; 0 other;
```

```
Query Match 63.1%; Score 1263; DB 24; Length 4825;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTTGACCGCTTATCAACAGCTTTAGAGAA 60
Db 1757 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTTGACCGCTTATCAACAGCTTTAGAGAA 1816
```

61 GTATTATCGTGGCTCTTTCTGTGAGGACACATGCAAGCAGAGAGATTTCTAAT 120
Db
1817 GTATTATCGTGGCTCTTTCTGTGAGGACACATGCAAGCAGAGAGATTTCTAAT 1876
Qy
121 GATGTGAAGTGGTGAAGACCAAGTTTCTACTCATGAGGGGTACATGATGATTTGACA 180
Db
1877 GATGTGAAGTGGTGAAGACCAAGTTTCTACTCATGAGGGGTACATGATGATTTGACA 1936
Qy
181 GCCCATCAGGCGCGGTGGTAAATATTCTACAAATGGGAAGTAAGCTGATTTGAAACAGGA 240
Db
1937 GCCCATCAGGCGCGGTGGTAAATATTCTACAAATGGGAAGTAAGCTGATTTGAAACAGGA 1996
Qy
241 AAATTATCAGAAGTGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 300
Db
1997 AAATTATCAGAAGTGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 2056
Qy
301 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db
2057 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 2116
Qy
361 GATCTCAGAAATC-GAAACTGAAAGTGAAGTGAATGACTGGCTAAACAAACAGAGAAAGA 419
Db
2117 GATCTCAGAAATCAGAAACTGAAAGTGAAGTGAATGACTGGCTAAACAAACAGAGAAAGA 2176
Qy
420 ACAAGGAATGAGGAGAGAGCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 479
Db
2177 ACAAGGAATGAGGAGAGAGCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTA 2236
Qy
480 CAACAAATAGGTGCTTCAAGAAAGTCTAGAACAAAGAACTCAAGTCAAGGTCAATTCCTC 539
Db
2237 CAACAAATAGGTGCTTCAAGAAAGTCTAGAACAAAGAACTCAAGTCAAGGTCAATTCCTC 2296
Qy
540 ACTCAATGCTGCTGATGATGATGATCTAGTGGAGATCAGCCTGCTGCTGTTGGA 599
Db
2297 ACTCAATGCTGCTGATGATGATGATCTAGTGGAGATCAGCCTGCTGCTGTTGGA 2356
Qy
600 GAACAACTTAAGTATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGC 659
Db
2357 GAACAACTTAAGTATTGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCGC 2416
Qy
660 TGGGTTCTTTTACAGACATCTCTTCAATGGCAAGCTCTTACTGAAAGAACAGTGCCTT 719
Db
2417 TGGGTTCTTTTACAGACATCTATAGATTACTGCAACAGTCTCCCTCGACCTGGAAAG 2476
Qy
720 TTTAGTCTAGGCTTTTACAGAAAGAGATGAGTGAACAGAGATTCACAACTGCTTT 779
Db
2477 TTTCTGCTGGCTTTACAGAGCTGAAACMACTGCCAATGTCTACAGATGCTACCGT 2536
Qy
780 A-----AAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACCTGGCCGTT 824
Db
2537 AAGGAAAGGCTCCTAGAGACTCCAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGAC 2596
Qy
825 TTAAGACCGGATCTAGAAAGAAAGACATCATGGGCAAACTGTATTCATCTCAACAA 884
Db
2597 CTCAGGGTGAATTTGAAGTCTACAGATGTTTATCAACAACCTGGATGAAACAGCCAA 2656
Qy
885 GATCTCTTTTCAACACTGAAGATAAGTCACTGACCCAGAACGGAAGCATGGCTGGAT 944
Db
2657 AAATCTGAGATCCCTGGAGGTTCCGATGATGAGTCTCTGTTACAAAGAGCTTTGGAT 2716
Qy
945 AACTTTGCCCGGTGGGTAATTTAGTCCAAAACTTTGAAAGAGTACAGCAC----- 999
Db
2717 AACATGAATCTCAAGTGGAGTGAACCTCGGAAAAAGTCTCTCAACATTAGTGGCCCATTTG 2776
Qy
1000 ----- 999
Db
2777 GAAGCCAGTTCTGACCAAGTGAAGGCTCTGCACCTTTCTCTCAGGAACTCTCTGGTGG 2836
Qy
1000 ----- 999
Db
2837 CTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGCACTATTGGAGGCGACTTTCCAGCA 2896
Qy
1000 ----- 999

2897 GTTCAGAAGCAGAACGATGTACATAGGCGCTTCAAGAGGGAATTTGAAAACTTAAAGACCT 2956
Qy
1000 ----- 999
Db
2957 GTAATCATGACTACTCTTGAGACTGTACGAATATTTCTGACAGAGACGCCCTTTTGAAGGA 3016
Qy
1000 ----- 999
Db
3017 CTAGAGAACTCTTACCAGGAGCCAGAGAGCTGCTCTCTGAGAGAGAGCCAGATGTC 3076
Qy
1000 ----- 999
Db
3077 ACTCGGCTTTACGAAAGCAGGCTGAGAGGCTCAATACTAGTGGGAAAAATTTGAAACCTG 3136
Qy
1000 -----AGACCCCTTGAAGACTCCAGGAACTTCAA 1028
Db
3137 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTGAAGACTCCAGGAACTTCAA 3196
Qy
1029 GAGGCCACGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGCATCAAGGGATCCTGG 1088
Db
3197 GAGGCCACGATGAGCTGGACCTCAAGCTGGCCCAAGCTGAGTGCATCAAGGGATCCTGG 3256
Qy
1089 CAGCCGCTGGCGGATCTCTCAATGACTCTCTCCAAAGTCACTTCGAGAAAGTCAAGGCA 1148
Db
3257 CAGCCGCTGGCGGATCTCTCAATGACTCTCTCCAAAGTCACTTCGAGAAAGTCAAGGCA 3316
Qy
1149 CTTGAGAGAGAAATTTGGCCCTCTGAAAGAGAACTGAGGCCACGTCATGACCTTGTCTGC 1208
Db
3317 CTTGAGAGAGAAATTTGGCCCTCTGAAAGAGAACTGAGGCCACGTCATGACCTTGTCTGC 3376
Qy
1209 CAGCTTACCCTTTGGGCACTTCAAGCTCTCAAGCTGATTAACCTCAGACTCTTGGAGAGACTG 1268
Db
3377 CAGCTTACCCTTTGGGCACTTCAAGCTCTCAAGCTGATTAACCTCAGACTCTTGGAGAGACTG 3436
Qy
1269 AACACCAGATGGAAGTCTCTCAGGCTGGCCGTGAGGACCGAGTCAGGCTCAGGAGTGCATGAA 1328
Db
3437 AACACCAGATGGAAGTCTCTCAGGCTGGCCGTGAGGACCGAGTCAGGCTCAGGAGTGCATGAA 3496
Qy
1329 GCCACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTGTCAGGGTCCC 1388
Db
3497 GCCACAGAGGACTTTGGTCCAGCATCTCAGCACTTTCTTCCAGCTCTGTCAGGGTCCC 3556
Qy
1389 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTATATATCAACACAGAGCTCAAAACA 1448
Db
3557 TGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTATATCAACACAGAGCTCAAAACA 3616
Qy
1449 ACTTGTGGGACACATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAT 1508
Db
3617 ACTTGTGGGACACATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAT 3676
Qy
1509 GTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGCTGCAAGAGGCCCTTTGC 1568
Db
3677 GTCAGATTTCTCAGCTTATAGGACTGCCATGAACTCCGAAAGCTGCAAGAGGCCCTTTGC 3736
Qy
1569 TTGGATCTTTGAGCCCTGTCAGCTGCAATGTGATGCCCTTGGACAGGACCAACCTCAAGCAA 1628
Db
3737 TTGGATCTTTGAGCCCTGTCAGCTGCAATGTGATGCCCTTGGACAGGACCAACCTCAAGCAA 3796
Qy
1629 AATGACAGCCCATGATATCTTGCAGATTTAAATGTTTGGACCACTATTTATGACCCG 1688
Db
3797 AATGACAGCCCATGATATCTTGCAGATTTAAATGTTTGGACCACTATTTATGACCCG 3856
Qy
1689 CTGAGCAAGAGACCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATATGTGTCTGAAC 1748
Db
3857 CTGAGCAAGAGACCAACAAATTTGGTCAAGCTCCCTCTCTGCGTGGATATGTGTCTGAAC 3916
Qy
1749 TGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGCTCTCTTTTAA 1808
Db
3917 TGGCTGCTGAATGTTTATGATACGGGACGAAACAGGAGGATCCGTGCTCTCTTTTAA 3976
Qy
1809 ACTGGCATCTTTCCCTGCTGTAAGAGCACTTTGGAAGACAAAGTACAGATACCTTTTCAAG 1868

3977 ACTGGCATTCTCCCTGTGTAAGACACATTTGGAAGACAAAGTACAGATACCTTTTCAAG 4036
1869 CAAGTGGCAAGTTCACAGGATTTTGTACAGCGCAGGCTGGGCTCTCTCATGAT 1928
4037 CAAGTGGCAAGTTCACAGGATTTTGTACAGCGCAGGCTGGGCTCTCTCATGAT 4096
1929 TCTATCCAAATTCACAGACAGTGGGTGAAGTTCATCTTTGGGGCGAGTAAACATTGAG 1988
4097 TCTATCCAAATTCACAGACAGTGGGTGAAGTTCATCTTTGGGGCGAGTAAACATTGAG 4156
1989 CCAAGTGTCCGGA 2001
4157 CCAAGTGTCCGGA 4169

RESULT 10
AAD37263
ID AAD37263 standard; DNA; 4848 BP.
AC AAD37263;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-CMV-delta3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Unidentified.
XX
PN WO200183695-A2.
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
PI Xiao X;
XX
DR WPI; 2002-049342/06.
XX

New dystrophin minigene for treating Duchenne or Becker muscular dystrophy comprises an N-terminal domain or modified N-terminal domain, rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin gene -
Example 1; Page 68-70; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding a dystrophin minigene. The minigene comprises N-terminal or modified N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4 domains and cysteine-rich domains of dystrophin or utrophin genes. The invention also relates to a recombinant adeno-associated virus (AAV) comprising dystrophin minigene operably linked to an expression control element. The dystrophin minigene in operable linkage with an expression control element, in a recombinant adeno-associated virus or retrovirus is useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) in a mammalian subject. The present sequence is AAV vector plasmid construct containing human dystrophin minigenes, a cytomegalovirus (CMV) promoter and a small polyA signal sequence.
XX
SQ Sequence 4848 BP; 1405 A; 1144 C; 1180 G; 1119 T; 0 other;

Query Match 63.1%; Score 1263; DB 24; Length 4848;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

QY 1 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 60
DB 1780 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 1839
QY 61 GTATATTCGTGGCTTCTCTGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAAT 120
DB 1840 GTATATTCGTGGCTTCTCTGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAAT 1899
QY 121 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 180
DB 1900 GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTGACA 1959
QY 181 GCCCATCAGGGCGGGTGGTAAATTTCTACAAATTTGGGAAGTAAAGCTGATTCGGAACAGA 240
DB 1960 GCCCATCAGGGCGGGTGGTAAATTTCTACAAATTTGGGAAGTAAAGCTGATTCGGAACAGA 2019
QY 241 AAATATCAGAAAGATGAAGAACTGAAGTACAAAGACAGATGAATCTCTCTAAATCAAGA 300
DB 2020 AAATATCAGAAAGATGAAGAACTGAAGTACAAAGACAGATGAATCTCTCTAAATCAAGA 2079
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAACCAAGCAATTTACATAGAGTTTAAATG 360
DB 2080 TGGGAATGCTCAGGGTAGCTAGCATGGAAACCAAGCAATTTACATAGAGTTTAAATG 2139
QY 361 GATCTCCAGATC-GAAACTGAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGA 419
DB 2140 GATCTCCAGATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAACAGAGAAAGA 2199
QY 420 ACAAGGAAATGGAGGAGAGCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 479
DB 2200 ACAAGGAAATGGAGGAGAGCTCTTGGACCTGATCTTTGAAGACCTAAACGCCAAGTA 2259
QY 480 CAACACATTAAGGTGCTTCAAGAGATCTAGAACAGACAACTAGTCAAGGTCAATTTCTCTC 539
DB 2260 CAACACATTAAGGTGCTTCAAGAGATCTAGAACAGACAACTAGTCAAGGTCAATTTCTCTC 2319
QY 540 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 599
DB 2320 ACTCACATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2379
QY 600 GAACAACTTAAGGTATGGAGATCGATGGGCAACATCTGTAGTGGACAGAGACAGCCG 659
DB 2380 GAACAACTTAAGGTATGGAGATCGATGGGCAACATCTGTAGTGGACAGAGACAGCCG 2439
QY 660 TGGGTTCCTTTTACAGACATCTCTCTCAATGGCAACCTGCTTACTGAAGAACAGTGCCTT 719
DB 2440 TGGGTTCCTTTTACAGACATCTCTCTCAATGGCAACCTGCTTACTGAAGAACAGTGCCTT 2499
QY 720 TTTAGTGCATGGCTTTTCAAGAAAGAGATGAGTGAAGCAAGATTCACACAACTGGCTTT 779
DB 2500 TTTCTTGGCTGGCTTACAGAGCTGAACCACTGCGCAATGTCTCTACAGGATGCTACCCGT 2559
QY 780 A-----AAGATCAAAATGAAATGTTATCAAGTCTTCAAAACCTGCGCGTT 824
DB 2560 AAGAAAGGCTCTAGAGACCTCAAGAGAGTAAAGAGCTGATGAACATGGCAAGAC 2619
QY 825 TTAAGAGCGGATCTAGAAAGAAAGAAAGCAATCCATGGGCAAACTGTATTCATCTCAAGAA 884
DB 2620 CTCAAGGTGAATTTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAACAGCCAA 2679
QY 885 GATCTCTTTTCAACCTGAAGATTAAGTCAGTGACCCAGAGAGGAGCATGGCTGGAT 944
DB 2680 AAATCTTGAGATCTCTGGAGAGTTCGGATGATGAGTCTCTGTACAAAGAGCTTTGGAT 2739
QY 945 AACTTTTGGCGGTGGGATTAATTTAGTCCAAAACTTTGAAAGAGTACAGCAC----- 999
DB 2740 AACATGAATTCAGTGGAGTGAACCTCGAAAAAGTCTCTCAACATTAGTCCCATTTG 2799
QY 1000 ----- 999
DB 2800 GAAGCCAGTTCTGACCAGTGGAGAGCGTCTGCACCTTTCTCTCAGGAACTTCTGTGTGG 2859
QY 1000 ----- 999

```
Db 2860 CTACAGCTGAAGATGATGAATTAAGCCGCGCAGCACCTATTAGGCGGACCTTTCCAGCA 2919
Qy 1000 ----- 999
Db 2920 GTTCAGAGCAGAACGATGTACATAGGGCTTCAAGAGGGAATTGAATACTAAGAACCT 2979
Qy 1000 ----- 999
Db 2980 GTAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCCTTTGGAAGCA 3039
Qy 1000 ----- 999
Db 3040 CTAGAGAAACTTACAGAGAACCCAGAGAGCTGCTCCTGAGGAGAGAGCCAGAGATGTC 3099
Qy 1000 ----- 999
Db 3100 ACTCGGCTTCTACGAAAGCAGGTGAGAGGTCATATCTAGTGGGAAAAAATTGAACCTG 3159
Qy 1000 -----AGACCTTGAAGACTCCAGGACTTCAA 1028
Db 3160 CACTCCGCTGACTGGCAGAGAAAAAATAGATGAGACCTTTGAAGAGCTCCAGGAACTTCAA 3219
Qy 1029 GAGGCCACGATGAGCTGGACCTCAAGCTCGGCAAGCTGAGGTGATCAAGGATCCTCG 1088
Db 3220 GAGGCCACGATGAGCTGGACCTCAAGCTCGGCAAGCTGAGGTGATCAAGGATCCTCG 3279
Qy 1089 CAGCCGCTGGCGGATCTCTCATATGACTCTCTCAAGATCCTCGAGAAAGTCAAGGCA 1148
Db 3280 CAGCCGCTGGCGGATCTCTCATATGACTCTCTCAAGATCCTCGAGAAAGTCAAGGCA 3339
Qy 1149 CTTTCGAGGAGAAATGGCCCTCTGAAGAAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCA 1208
Db 3340 CTTTCGAGGAGAAATGGCCCTCTGAAGAAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCA 3399
Qy 1209 CAGCTTACCACTTTGGCAATTCAGCTCTCAAGTATTAACCTCAGCACTCTGGAAGACCTG 1268
Db 3400 CAGCTTACCACTTTGGCAATTCAGCTCTCAAGTATTAACCTCAGCACTCTGGAAGACCTG 3459
Qy 1269 AACACGATGGAAGCTTTCGAGGTGGCGTTCGAGGACCGAGTCAAGGAGTCAAGTGA 1328
Db 3460 AACACGATGGAAGCTTTCGAGGTGGCGTTCGAGGACCGAGTCAAGGAGTCAAGTGA 3519
Qy 1329 GCCACAGGAGCTTTGGTCCAGATCTCAGCACTTTCTTCCACGCTCTGTCAGGCTCC 1388
Db 3520 GCCACAGGAGCTTTGGTCCAGATCTCAGCACTTTCTTCCACGCTCTGTCAGGCTCC 3579
Qy 1389 TGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTATATATCAACCAAGAGACTCAACA 1448
Db 3580 TGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTATATATCAACCAAGAGACTCAACA 3639
Qy 1449 ACTTGTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGTCACTGAATAAT 1508
Db 3640 ACTTGTGGGACCATCCAAAATGACAGAGCTCTACAGTCTTTAGTCACTGAATAAT 3699
Qy 1509 GTCAGATTCTCAGCTTATAGAGCTGCCATGAATCCGAAAGACTGCGAGAGGCGCTTTC 1568
Db 3700 GTCAGATTCTCAGCTTATAGAGCTGCCATGAATCCGAAAGACTGCGAGAGGCGCTTTC 3759
Qy 1569 TTGATCTCTTACGCTTGTGATGTGATGCTTGAACAGAGCACTCAAGCA 1628
Db 3760 TTGATCTCTTACGCTTGTGATGTGATGCTTGAACAGAGCACTCAAGCA 3819
Qy 1629 AATCAGCAGCCATGGATATCTCTCAGATTATTAATTTGACCACTATTTATGACCG 1688
Db 3820 AATCAGCAGCCATGGATATCTCTCAGATTATTAATTTGACCACTATTTATGACCG 3879
Qy 1689 CTGAGCAGAGAGCAACAATTTGGTCAAGCTCCCTCTCTGGTGGATATGTCTGAAC 1748
Db 3880 CTGAGCAGAGAGCAACAATTTGGTCAAGCTCCCTCTCTGGTGGATATGTCTGAAC 3939
Qy 1749 TGGCTGCTGAATGTTATGATACGGGAGCAACAGGAGGATCCGTGCTCTTTTAAA 1808
```

```
Db 3940 TGGCTGCTGAATGTTATGATACGGGAGCAACAGGAGGATCCGTGCTCTTTTAAA 3999
Qy 1809 ACTGGCATATTTCCCTGTGTAAAGCACATTTGAGACAGTACAGATACCTTTTCAAG 1868
Db 4000 ACTGGCATATTTCCCTGTGTAAAGCACATTTGAGACAGTACAGATACCTTTTCAAG 4059
Qy 1869 CAAGTGGCAAGTTCAAACAGGATTTGTGACAGCGCAGGCTGGCCCTCTCTCATGAT 1928
Db 4060 CAAGTGGCAAGTTCAAACAGGATTTGTGACAGCGCAGGCTGGCCCTCTCTCATGAT 4119
Qy 1929 TCTATCCAAATTCAAAGACAGTTCGGTGAAGTTGCATCTTTTGGGGCAGTAACATTCAG 1988
Db 4120 TCTATCCAAATTCAAAGACAGTTCGGTGAAGTTGCATCTTTTGGGGCAGTAACATTCAG 4179
Qy 1989 CCAAGTGTCCGA 2001
Db 4180 CCAAGTGTCCGA 4192
```

RESULT 11
AAD37264
ID AAD37264 standard; DNA; 5060 BP.
XX
AC AAD37264;
XX
DT 21-AUG-2002 (first entry)
XX
DE Adeno-associated virus (AAV) vector plasmid, AAV-B-CMV-3849.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Cytomegalovirus.
XX Chimeric - Unidentified.
PN WO200183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX (XIAO/) XIAO X.
XX Xiao X;
XX
XX WPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene -
XX
PS Example 1; Page 70-71; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
CC vector plasmid construct containing human dystrophin minigenes, a
CC muscle creatine kinase (MCK) enhancer, a cytomegalovirus (CMV) promoter
CC and a small polyA signal sequence.
XX
SQ Sequence 5060 BP; 1449 A; 1217 C; 1234 G; 1160 T; 0 other;

Query Match 63.1%; Score 1263; DB 24; Length 5060;			
Best Local Similarity 75.3%; Pred. No. 0; Mismatches 185; Indels 412; Gaps 3;			
Matches 1816; Conservative 0;			
QY	1	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA	60
DB	1992	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA	2051
QY	61	GTATTAATCGTGGCTTTCTCTGCTGAGGACATTTGCAAGCAGACAGGAGAGATTCTTAAT	120
DB	2052	GTATTAATCGTGGCTTTCTCTGCTGAGGACATTTGCAAGCAGACAGGAGAGATTCTTAAT	2111
QY	121	GATGTGGAAGTGGTGAAGACAGCTTTATCATCTCATGAGGGGTACATGATGGAATTGACA	180
DB	2112	GATGTGGAAGTGGTGAAGACAGCTTTATCATCTCATGAGGGGTACATGATGGAATTGACA	2171
QY	181	GCCATCAGGGCGGGTGGTGAATATTTACAAATTTGGGGAAGTGAAGTGAACAGGA	240
DB	2172	GCCATCAGGGCGGGTGGTGAATATTTACAAATTTGGGGAAGTGAAGTGAACAGGA	2231
QY	241	AAATTAATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGA	300
DB	2232	AAATTAATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAGA	2291
QY	301	TGGGAATGCCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG	360
DB	2292	TGGGAATGCCCTCAGGGTAGCTAGCATGGAAACAAAGCAATTTACATAGAGTTTAAATG	2351
QY	361	GATCTCCAGATC-GAACTGAAGAGTTGAATGACTGGCTAACAAACAAACAGAGAGAA	419
DB	2352	GATCTCCAGATCAGAACTGAAGAGTTGAATGACTGGCTAACAAACAAACAGAGAGAA	2411
QY	420	ACAAGAAATGGAGAGAGGCTCTTTGACCTGATCTTTGAAGACTTAAACGCCAAGTA	479
DB	2412	ACAAGAAATGGAGAGAGGCTCTTTGACCTGATCTTTGAAGACTTAAACGCCAAGTA	2471
QY	480	CAACACATAGGTGCTTCAAGAGATCTAGAACAAAGACAGTCAAGGTCAATTTCTCTC	539
DB	2472	CAACACATAGGTGCTTCAAGAGATCTAGAACAAAGACAGTCAAGGTCAATTTCTCTC	2531
QY	540	ACTCATATGCTGGTGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGAA	599
DB	2532	ACTCATATGCTGGTGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTGAA	2591
QY	600	GAACAACTTAAGTATTTGGAGATCATGGGCAACATCTGTAGATGAACAGAACCCGC	659
DB	2592	GAACAACTTAAGTATTTGGAGATCATGGGCAACATCTGTAGATGAACAGAACCCGC	2651
QY	660	TGGGTTCTTTTACAAGACATCTCTTCAATGGCAAGCTTTACTGAAGACAGTGCCTT	719
DB	2652	TGGGTTCTTTTACAAGACATCTCTTCAATGGCAAGCTTTACTGAAGACAGTGCCTT	2711
QY	720	TTTATGTGATGGCTTTTCAAGAAAGAGATGCGATGAACAGATTCACACAACTGGCTTT	779
DB	2712	TTTCTTGGCTGGCTTACAGAACTGAACAACTGCGCAATGTCTCAGGATGCTACCCGT	2771
QY	780	A-----AAGATCAAAATGAATGTATCAAGTCTTCAAACTGCAAACTGGCCGT	824
DB	2772	ARGGAAAGGCTCTAGAAAGCTCCAGGGAGTAAAGAGCTGATGAACAAATGGCAAGAC	2831
QY	825	TTAAAGCGGATCTAGAAAGAAAGCAATCTGGGCAAACTGATTTCACTCAAAACAA	884
DB	2832	CTCCAGGTGAATTTGAGCTCACAGATGTTTATCAACAACTGGATGAAGACAGCAA	2891
QY	885	GATCTTTTCAACATCAAGATTAAGTCAAGTCAAGCCAGAACAGGAGCATGGCTGAT	944
DB	2892	AAAATCTGAGATCCCTGGAAGTTCGATGATGCAATCTGTTTCAAAAGACGTTGGAT	2951
QY	945	AACCTTTCCCGGTGTTGGGATAATTTAGTCCAAACAACTTGAAGAGTACAGCAC-----	999
DB	2952	AACATGAACCTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTAGTCCCATTTG	3011
QY	1000	-----	999
DB	3012	GAAGCCAGTTCTGACCAGTGGAGAGCTCTGCACCTTTCTCTGCAGGAACCTTCTGGTGTG	3071
QY	1000	-----	999
DB	3072	CTACAGCTGAAGATGATGAATTAAGCCGAGGACCTATTGGAGGCGACTTTCCAGCA	3131
QY	1000	-----	999
DB	3132	GTTCAAGCAGAACGATGTACATAGGGCTTCAAGGGGAATTGAAACATAAGAACCT	3191
QY	1000	-----	999
DB	3192	GTAATCATGAGTACTCTTGAGACTGTACGAATTTCTGCACAGCAGCCTTTGGAAGGA	3251
QY	1000	-----	999
DB	3252	CTAGAGAACTCTACAGGAGCCAGAGAGCTGCCTCTGAGGAGAGAGCCAGAAATGTC	3311
QY	1000	-----	999
DB	3312	ACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATCTAGTGGGAAAAAATTGAACCTG	3371
QY	1000	-----AGACCTTTGAAGACTCCAGGAACCTCCAGGAACCTTCAA	1028
DB	3372	CATCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAGACTCCAGGAACCTCAA	3431
QY	1029	GAGCCACGATGAGCTGGACCTCAAGCTCGGCAAGCTGAGGTGATCAAGGGATCTCTCG	1088
DB	3432	GAGCCACGATGAGCTGGACCTCAAGCTCGGCAAGCTGAGGTGATCAAGGGATCTCTCG	3491
QY	1089	CAGCCGCTGGGCGATCTCTCATTTGATCTCTTCAAGATCACTCGAGAACTCAAGGCA	1148
DB	3492	CAGCCGCTGGGCGATCTCTCATTTGATCTCTTCAAGATCACTCGAGAACTCAAGGCA	3551
QY	1149	CTTCGAGGAAATTCGGCTCTGAAAGAAAGCTGAGCCAGCTCAATGACCTTTGCTCGC	1208
DB	3552	CTTCGAGGAAATTCGGCTCTGAAAGAAAGCTGAGCCAGCTCAATGACCTTTGCTCGC	3611
QY	1209	CAGCTTACCACTTTGGGCAATCAGCTCTCACGCTAATGAGTAACTGACCTTTGCTCGC	1268
DB	3612	CAGCTTACCACTTTGGGCAATCAGCTCTCACGCTAATGAGTAACTGACCTTTGCTCGC	3671
QY	1269	AACACAGATGGAAGCTTTCTGAGGTGGCGTGGAGACCGAGTCAAGGAGCTGCAAGAA	1328
DB	3672	AACACAGATGGAAGCTTTCTGAGGTGGCGTGGAGACCGAGTCAAGGAGCTGCAAGAA	3731
QY	1329	GCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGCTCTGTCCAGGTC	1388
DB	3732	GCCACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTTCCACGCTCTGTCCAGGTC	3791
QY	1389	TGGAGAGACCACTCTGGCCAAACAAAGTCCCTACTATCAACACAGAGACTCAAAACA	1448
DB	3792	TGGAGAGAGCCATCTCGCCAAACAAAGTCCCTACTATCAACACAGAGACTCAAAACA	3851
QY	1449	ACTTGTCTGGACCACTCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCTGGAATA	1508
DB	3852	ACTTGTCTGGACCACTCCCAAAATGACAGAGCTCTACCACTTTTAGCTGACCTGGAATA	3911
QY	1509	GTCAATCTCAGCTTATAGGATGCGCATGAAACTCCGAGAGTCAAGAGGCGCTTTGC	1568
DB	3912	GTCAATCTCAGCTTATAGGATGCGCATGAAACTCCGAGAGTCAAGAGGCGCTTTGC	3971
QY	1569	TTGGATCTCTTGAAGCTGTCAGTGTGATGTCCTTGGACAGACCAACCTTCAAGCAA	1628
DB	3972	TTGGATCTCTTGAAGCTGTCAGTGTGATGTCCTTGGACAGACCAACCTTCAAGCAA	4031
QY	1629	AATCAACAGCCCAATGGATATCTTCAGAGATTTAAATTTGTTGACCACTATTTATGACCC	1688
DB	4032	AATCAACAGCCCAATGGATATCTTCAGAGATTTAAATTTGTTGACCACTATTTATGACCC	4091
QY	1689	CTGAGCAAGAGCAGACAAATTTGGTCAACGCTCTCTCTGCGTGGATATGTCTCTGAAC	1748

Db 4092 CTGGAGCAAGCAGACCAATTTGGTCAACGTCCTCTCTGGGTGATATGTCCTGAAC 4151
Qy 1749 TGGCTGCTGAATTTATGATACGGGACGAACAGGAGGATCCGTCTCTGCTTTTAAA 1808
Db 4152 TGGCTGCTGAATTTATGATACGGGACGAACAGGAGGATCCGTCTCTGCTTTTAAA 4211
Qy 1809 ACTGGCATATTTCCCTGTGTAAAGACATTTGGAAAGACAAAGTACAGATACCTTTTCAAG 1868
Db 4212 ACTGGCATATTTCCCTGTGTAAAGACATTTGGAAAGACAAAGTACAGATACCTTTTCAAG 4271
Qy 1869 CAAGTGCAGTTCAACAGGATTTGTGCCAGCGAGGCTGGCTCTCTGCGATGAT 1928
Db 4272 CAAGTGCAGTTCAACAGGATTTGTGCCAGCGAGGCTGGCTCTCTGCGATGAT 4331
Qy 1929 TCTATCCAAATTCGAACAGATTTGGGTGAAGTTGCAATCCTTTGGGGGCGAGTAAACATTGAG 1988
Db 4332 TCTATCCAAATTCGAACAGATTTGGGTGAAGTTGCAATCCTTTGGGGGCGAGTAAACATTGAG 4391
Qy 1989 CCAAGTGTCCGA 2001
Db 4392 CCAAGTGTCCGA 4404

RESULT 12
ABK81998
ID ABK81998 standard; DNA; 5339 BP.
AC ABK81998;
XX
XX
XX 13-AUG-2002 (first entry)
XX
XX DNA encoding mini-dystrophin protein deltaR2-R21.
XX
XX Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
XX Duchenne's muscular dystrophy; DMD; dystrophin; ds.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX W0200229056-A2.
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-US31126.
XX
XX 06-OCT-2000; 2000US-238848P.
XX
XX (UNMI) UNIV MICHIGAN.
XX
XX Chamberlain JS, Harper SQ;
XX
XX WPI; 2002-435334/45.
XX

XX
XX A composition for preparing therapeutic drugs, has a mini-dystrophin
XX peptide comprising a specific number of spectrin-like repeat domains,
XX or a nucleic acid sequence encoding the mini-dystrophin peptide -
XX
XX Example 6; Fig 13; 145pp; English.
XX
XX The invention describes a composition comprising a mini-dystrophin
XX peptide comprising a spectrin-like repeat domain, where the domain
XX comprises n spectrin-like repeats, and contains no more than n
XX spectrin-like repeats, where n is an even number between 4-24, or a
XX nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
XX peptide or the polynucleotide encoding it is useful as a medicament
XX for preparing a drug for therapeutic application and in the preparation
XX of a composition for treatment of muscle disease, e.g. Duchenne's
XX muscular dystrophy (DMD). This sequence represents a mini-dystrophin
XX sequence of the invention.

XX Sequence 5339 BP; 1638 A; 1191 C; 1187 G; 1323 T; 0 other;

Query Match 62.7%; Score 1254; DB 24; Length 5339;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1621; Conservative 0; Mismatches 380; Indels 79; Gaps 6;
Qy 1 GGCAGTTTCATTGTGAGAGTGAAGTAAACCTCGACCGTTATCAACAGCTTTAGAGAA 60
Db 1199 GGCAGTTTCATTGTGAGAGTGAAGTAAACCTCGACCGTTATCAACAGCTTTAGAGAA 1258
Qy 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAAT 120
Db 1259 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTCTTAAT 1318
Qy 121 GATGTGAAGTGGTGAAGACCCAGTTTCATCTACTCATGAGGGTACATGATGGATTTCACA 180
Db 1319 GATGTGAAGTGGTGAAGACCCAGTTTCATCTACTCATGAGGGTACATGATGGATTTCACA 1378
Qy 181 GCCCATCAGGCGCCGGTGGTAAATTTCTACAATTGGGAAAGTAAAGCTGATTGGAACAGGA 240
Db 1379 GCCCATCAGGCGCCGGTGGTAAATTTCTACAATTGGGAAAGTAAAGCTGATTGGAACAGGA 1438
Qy 241 AAATTATCAGAAGATGAAGAAACTGAAAGTACAGAGCAGATGAATCTCCTAAATTCAAGA 300
Db 1439 AAATTATCAGAAGATGAAGAAACTGAAAGTACAGAGCAGATGAATCTCCTAAATTCAAGA 1498
Qy 301 TGGGAATGCCCTCAGGCTAGCTAGCATGCGAAACCAAAAGCAATTTACATAGATTTTAAATG 360
Db 1499 TGGGAATGCCCTCAGGCTAGCTAGCATGCGAAACCAAAAGCAATTTACATAGATTTTAAATG 1558
Qy 361 GATCTCCAGAAATC-GAAACTGAAAGATTTGAATGATGCTGCTAAACAAACAGAGAAAGA 419
Db 1559 CAACAGTTCCCTCGGACCTGGAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACT 1618
Qy 420 ACAAGGAAATGGAGGAGAGGCTCTTGGACCTGATCTTGAAGACCTTAAACCCAGTA 479
Db 1619 GCCAATGTCTTACAGGATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTA 1678
Qy 480 CAACAAACATAGGTGCTTCAAGAGATCTAGAAACAGAAACAAGTCAGGGTCAATTTCTC 539
Db 1679 AAAGAGCTGATGAACAATGGCAAGACCTCCAAGTGAATTTGAAGCTCACACAGATGTT 1738
Qy 540 ACTCAGATGGTGGTGGTAG-----TTGATGAATCTAGTGGAGATCAGCAACT 587
Db 1739 TATCACAACCTGGATGAAACAGCCAAAAAATCTGAGATCCCTGGAAGGTTCCGATGAT 1798
Qy 588 GCTGCTTTTGAAGAAACAACTTAAAG---TATTGGAGATCGATGGGCAACATCTGTAGA 644
Db 1799 GCAGTCTGTTCAGAGACGTTGGATACATGAACTTCAAGTGGAGTGAATTCGGA 1858
Qy 645 TGGACAGAAGACCGCTGGGTTCTTTTACAGACATCCTTCTCAATGGCAACGCTTACT 704
Db 1859 AAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGAAGCGTCTGCAC 1918
Qy 705 GAAGAACAGTGCCTTTTGTGATGGCTTTTCAGAAAAAGAGATGCAAGTGAACAAGATT 764
Db 1919 CTTTCTCTGAGGAACTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCCGCGAG 1978
Qy 765 CACAAACTGGCTTTAAAGATCAAAATGAATGTATCAAGTCTTCAAAAACCTGGCGTT 824
Db 1979 GCACCTATTGAGGCGACTTTCAGAGATTTCCAGAGATTCAGAACCAAGCTGTACATAGGCGCTC 2038
Qy 825 TTAAAGCGGATCTAGAAAGAAAGCAATCCATGGGCAACTGTA----- 871
Db 2039 AAGAGGGAATTTGAAGAACTAAAGAACCTGTAAATCATGATGACTCTTGAGACTGTACGAATA 2098
Qy 872 -TTCACCTCAACAAAGATCTTCTTCAACACTGAAGAAATAGTCAGTGAACCCAGAGACGG 930
Db 2099 TTTCTGACAGAGAGCCCTTTGGAGGACTAGAGAACTCTACAGAGGCCACAGAGCTG 2158
Qy 931 AAGCATGGCTGGA-----TAACTTTGCCCGTGTGGG 963
Db 2159 CCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGTC 2218
Qy 964 ATAATTAGTCCAAAAACTTGAA-----AAGACTACAGCAG 1001

Db 2219 AATACCTAGTGGGAAAATTTGAACCTGCACTCCGCTGACTGCGACGAGAAAATAGATGAG 2278
QY 1002 ACCCTTGAAGACTCCAGAACTTCAAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGC 1061
Db 2279 ACCCTTGAAGACTCCAGAACTTCAAGAGGCCACGGATGAGCTGAGCTCAAGCTGCGC 2338
QY 1062 CAAGCTGAGTATCAAGGATCTCTGGCAGCCGCTGGGGATCTCTCATTTGACTCTCTC 1121
Db 2339 CAAGCTGAGTATCAAGGATCTCTGGCAGCCGCTGGGGATCTCTCATTTGACTCTCTC 2398
QY 1122 CAAGATCACTCGAGAACTCAAGGACTTTCAGAGAGAAATTTGGCCCTCTGAAGAGAAC 1181
Db 2399 CAAGATCACTCGAGAACTCAAGGACTTTCAGAGAGAAATTTGGCCCTCTGAAGAGAAC 2458
QY 1182 GTGAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCAACG 1241
Db 2459 GTGAGCCAGCTCAATGACCTTGTCTGCCAGCTTACCACCTTTGGGCATTCAGCTCTCAACG 2518
QY 1242 TATACTCTCAGCACTCTGGAAGACCTTGAACACAGATGGAAGCTTTCTGAGGTGCCGCTC 1301
Db 2519 TATACTCTCAGCACTCTGGAAGACCTTGAACACAGATGGAAGCTTTCTGAGGTGCCGCTC 2578
QY 1302 GAGGACCGAGTCAAGGAGCTGATGAGGCCACAGGACTTTTGGTCCAGCATCTCAGCAC 1361
Db 2579 GAGGACCGAGTCAAGGAGCTGATGAGGCCACAGGACTTTTGGTCCAGCATCTCAGCAC 2638
QY 1362 TTTCTTTTCCAGCTCTCTCAGGCTCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCC 1421
Db 2639 TTTCTTTTCCAGCTCTCTCAGGCTCTCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCC 2698
QY 1422 TACTATATCAACACAGAGACTCAAACTTGTGGGACCATCCCAAAATGACAGACTC 1481
Db 2699 TACTATATCAACACAGAGACTCAAACTTGTGGGACCATCCCAAAATGACAGACTC 2758
QY 1482 TACCACTTTTACCTGACCTGATATATGTCAGATCTATAGGACTCCCATGAAA 1541
Db 2759 TACCACTTTTACCTGACCTGATATATGTCAGATCTATAGGACTCCCATGAAA 2818
QY 1542 CTCGAGAGACTGAGAGGCCCTTTCTTGGATCTCTTCCAGCTGCTGACCTGCATGTAT 1601
Db 2819 CTCGAGAGACTGAGAGGCCCTTTCTTGGATCTCTTCCAGCTGCTGACCTGCATGTAT 2878
QY 1602 GCCTTGGACAGCACACCTCAAGCAAAATGACCCATGATATCTCTGAGATTTT 1661
Db 2879 GCCTTGGACAGCACACCTCAAGCAAAATGACCCATGATATCTCTGAGATTTT 2938
QY 1662 AATTGTTTGACCACTATTTATGACCCGCTGGAGCAAGACCAAAATTTGTCACGTC 1721
Db 2939 AATTGTTTGACCACTATTTATGACCCGCTGGAGCAAGACCAAAATTTGTCACGTC 2998
QY 1722 CCTCTCTGCTGATATGCTCTGAATGCTGCTGCTGAATGTTTATGATACGGGACGACA 1781
Db 2999 CCTCTCTGCTGATATGCTCTGAATGCTGCTGCTGAATGTTTATGATACGGGACGACA 3058
QY 1782 GCGAGGATCCGTCTCTCTTTTAAACCTGGCATATTTCCCTGCTGTTAAAGCAATTTG 1841
Db 3059 GCGAGGATCCGTCTCTCTTTTAAACCTGGCATATTTCCCTGCTGTTAAAGCAATTTG 3118
QY 1842 GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAG 1901
Db 3119 GAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAG 3178
QY 1902 CGCAGGCTGGGCTCTCTCTGATGCTTATCAAAATTCAGAGAGTGGTGGTGAAGTT 1961
Db 3179 CGCAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTCAGAGAGTGGTGGTGAAGTT 3238
QY 1962 GCATCTCTTGGGGSCAGTAACTTGAAGCCCAAGTGTCCGGA 2001
Db 3239 GCATCTCTTGGGGSCAGTAACTTGAAGCCCAAGTGTCCGGA 3278

RESULT 13

AAAD37238
ID AAD37238 standard; DNA; 3531 BP.
XX
AC AAD37238;
XX
DT 21-AUG-2002 (first entry)
XX
DE Human dystrophin minigene delta3531.
XX
KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
KW Becker muscular dystrophy; ds.
XX
OS Homo sapiens.
XX
PN WO2001183695-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US13677.
XX
PR 28-APR-2000; 2000US-200777P.
XX
PA (XIAO/) XIAO X.
XX
PI Xiao X;
XX
XX MPI; 2002-049342/06.
XX
PT New dystrophin minigene for treating Duchenne or Becker muscular
PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
PT rod repeats, H1 and H4 domains and a cysteine rich domain of a
PT dystrophin gene.
XX
PS Example 1; Page 50-51; 71pp; English.
XX
CC The present invention relates to an isolated nucleotide sequence encoding
CC a dystrophin minigene. The minigene comprises N-terminal or modified
CC N-terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
CC invention also relates to a recombinant adeno-associated virus (AAV)
CC comprising dystrophin minigene operably linked to an expression control
CC element. The dystrophin minigene in operable linkage with an expression
CC control element, in a recombinant adeno-associated virus or retrovirus is
CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
CC dystrophy (BMD) in a mammalian subject. The present sequence is human
CC dystrophin minigene delta3531 containing nucleotides 1-1341 (N-terminus,
CC hinge H1 and rod R1), 8059-10227 (rods R22, R23 and R24, hinge H4 and
CC CR domain) and 11047-11058 (dystrophin last 3 amino acids).
XX
SQ Sequence 3531 BP; 1071 A; 809 C; 824 G; 827 T; 0 other;
Query Match 62.58; Score 1251.2; DB 24; Length 3531;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;
QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGACCGCTTATCAACAGCTTTAGAGAA 60
Db 1000 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGACCGCTTATCAACAGCTTTAGAGAA 1059
QY 61 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGACCAAGGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTTCTTTCTGCTGAGGACACATTCGACCAAGGAGAGATTTCTAAT 1119
QY 121 GATGTGAAGTGGTGAAGACCACTTTCTACTCATGAGGGGTACATGATGGATTTGACA 180
Db 1120 GATGTGAAGTGGTGAAGACCACTTTCTACTCATGAGGGGTACATGATGGATTTGACA 1179
QY 181 GCCCATCAGGCGCGGTGGTATATTTCTACATTTGGGAAGTGAAGTGAAGCAAGGA 240
Db 1180 GCCCATCAGGCGCGGTGGTATATTTCTACATTTGGGAAGTGAAGTGAAGCAAGGA 1339
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAAAGAGCAGATGAATCTCTCTAAATTTCAAGA 300

1240 AAATTATCAAGAGATGAGAACTGAAGTACAGAGCAGATGAATCTCTAAATTCAAGA 1299
301 TGGGAATGCCCTAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAGTTTAAATG 360
1300 TGGGAATGCCCTAGGGTAGCTAGCATGGAAAAACAAGCAATTTACATAGAACTCATAGA 1359
361 GATCTCCAGAA-----TCGAAACTGAAAGAGTTGAATGACTGGCTAAACAAAACAGAA 413
1360 TTACTGCAACAGTTCCCTCGACCTGGAAAGATTCTTGGCTGGCTTACAGAGCTGAA 1419
414 GAAAGAACAAAGGAAATGGAGGAGAGCCCTTTGGACCTGATCTTGAAGACCTTAAACGC 473
1420 ACAACTGCCAATGCTCTACAGGATGCTACCGTAAGGAAAGGCTCTAGAAAGACTCCAG 1479
474 CAAGTACAACAATAGGTCCTTCAGAGAGACTAGAACAGAAACAAGTCAAGGTCAT 533
1480 GGAGTAAAGAGCTGATGAACAATGCAAGACCTCCAGGTGAATGAAGTCAACA 1539
534 TCTCTCACTCACATGGTGGTGGTAGTTGATGA-----ATCTAGTGGAGATCAC 581
1540 GATGTTTATCACAACTGGATGAAACAGCCAAAATCTCGAGATCCCTGGAAGTTCC 1599
582 GCACTGCTCTTTGGAGACAACTTAAG-----TATGGGAGATCGATGGGCAACATC 638
1600 GATGATGCACTCCTGTACAAAGACGTTTGGATAACAATGAACTTCAAGTGGAGTGAAT 1659
639 TGTAGTGGACAGAGACCCCTGGGTTCTTTTACAAGACATCTTCTCAATGGCAACGT 698
1660 CGGAAAGTCTCTCAACATTAAGTCCCATTTGGAGCCAGTTCTCACCCAGTGGAGCGT 1719
699 CTCTACTGAAGAACAGTGCCTTTTATGTCAGTGGCTTTTCAAGAAAGAAAGATGAGTGAAC 758
1720 CTGCACCTTTCTCGAGGAACCTCTGGTGTGCTACAGCTGAAGATGATGAATTAAGC 1779
759 AAGATTACACAACTGCTTTTAAAGATCAAAATGAATGTATCAAGTCTTCAAAACTG 818
1780 CGGAGGCACTATTGGAGCGCATTTTCCAGCAGTTTCAGAGCAGTTCAGAGCAGAACGATGTACATAG 1839
819 GCCGTTTTTAAAGCGGATCTAGAAAAAGAAAGCAATCCATGGGCAAACTGTA----- 871
1840 GCCTTCAAGAGGAAATTGAACCTTAAAGAACCTGTATCATGATGACTCTCTTGAGACTGTA 1899
872 -----TTCACTCAACAGATCTTCTTTCACACTGAAGATAAGTCAAGTCAAGCCAGA 924
1900 CGAATATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAACTCTACCAAGGAGCCAGA 1959
925 AGCGGAAGCATGGCTGGA-----TAACTTTGCCCGGT 957
1960 GAGTGGCTCTGAGGAGAGGCCAGAAATGCTCACTCGGCTTCTAGAAAGCAGGCTGAG 2019
958 GTTGGGATAATTTAGTCCAAAACCTTGAA-----AAGATACA 995
2020 GAGTCAATATCTAGTGGGAAAAATTTGAACCTCGACTCGCTGACTGGCAGAGAAAAATA 2079
996 GCACAGACCTTTGAAGACTCCAGGACTTCAAGAGCCACGGATGAGTGGACCTCAAG 1055
2080 GATGAGACCTTTGAAGACTCCAGGAACTTCAAGAGCCACGGATGAGCTGGACCTCAAG 2139
1056 CTGCGCCCAAGCTGAGGTGATCAAGGGATCTCGGAGCCCGTGGCGGATCTCTCAATTGAC 1115
2140 CTGCGCAAGCTGAGGTGATCAAGGGATCTGGCAGCCCGTGGCGGATCTCTCAATTGAC 2199
1116 TCTCTCAAGATCACTCGAAGAAAGTCAAGGCACTTCAGAGAGAAATTCGCGCTCTGAA 1175
2200 TCTCTCAAGATCACTCGAAGAAAGTCAAGGCACTTCAGAGAGAAATTCGCGCTCTGAA 2259
1176 GAGAACGTGAGCCACGCTCAATGACTTGTCTCGCAGCTTACCACTTTGGGCATTCAGCTC 1235
2260 GAGAACGTGAGCCACGCTCAATGACTTGTCTCGCAGCTTACCACTTTGGGCATTCAGCTC 2319
1236 TCACCGTATAACCTCAGACCTCTCGAAGACCTGAAACCAAGATGGAAGCTTTGCGAGGTG 1295

2320 TCACCGTATAACCTCAGACCTCTGGAAGACTGAACACCAGATGGAAGCTTTCTCGAGGTG 2379
1296 GCCCTCGAGGACCCAGTCAAGCTCAGTCAAGAGGACTTTTGGTCCAGCATCT 1355
2380 GCCGTGAGGACCCAGTCAAGCTCAGTCAAGAGGACTTTTGGTCCAGCATCT 2439
1356 CAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAA 1415
2440 CAGCACTTTCTTTCCAGCTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAA 2499
1416 GTGCCCTACTATATCAACCAAGAGACTCAACACTTGTCTGGGACCATCCCAATATGACA 1475
2500 GTGCCCTACTATATCAACCAAGAGACTCAACACTTGTCTGGGACCATCCCAATATGACA 2559
1476 GAGCTCTACCAAGCTTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCC 1535
2560 GAGCTCTACCAAGCTTTTAGCTGACCTGAATAATGTCAGATTCTCAGCTTATAGACTGCC 2619
1536 ATGAACTCCGAAGACTGCGAAGGCCCTTTGCTTGAATCTCTTGAGCCTGTCAAGTGA 1595
2620 ATGAACTCCGAAGACTGCGAAGGCCCTTTGCTTGAATCTCTTGAGCCTGTCAAGTGA 2679
1596 TGTGATGCTTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG 1655
2680 TGTGATGCTTTGGACCAAGCAACCTCAAGCAAAATGACCAAGCCCATGGATATCCTGCAG 2739
1656 ATTATTAATGTTTGAACACTTATTTATGACCGCTGAGCAAGAGCAACAATTTGGTC 1715
2740 ATTATTAATGTTTGAACACTTATTTATGACCGCTGAGCAAGAGCAACAATTTGGTC 2799
1716 AACCTCCCTCTCTCGGTGGATATGCTGAACCTGCTGAATGTTTATGATACGGGA 1775
2800 AACGTCCCTCTCTCGGTGGATATGCTGAACCTGCTGAATGTTTATGATACGGGA 2859
1776 CGAACAGGGAGATCCGCTGCTCTTTTAAACTGGCATCAATTTCCCTGTGTAAAGCA 1835
2860 CGAACAGGGAGATCCGCTGCTCTTTTAAACTGGCATCAATTTCCCTGTGTAAAGCA 2919
1836 CATTTGGNAGACAGTACAGATACCTTTTCAAGCAAGTGGCAGTTTCAACAGGATTTGT 1895
2920 CATTTGGNAGACAGTACAGATACCTTTTCAAGCAAGTGGCAGTTTCAACAGGATTTGT 2979
1896 GACCAAGCAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGGT 1955
2980 GACCAAGCAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCAAAGACAGTTGGGT 3039
1956 GAAGTTCATCTTTGGGGGAGTAAACATTGAGCCAAAGTGTCCGGA 2001
3040 GAAGTTCATCTTTGGGGGAGTAAACATTGAGCCAAAGTGTCCGGA 3085

RESULT 14

AAD37258
ID AAD37258 standard; DNA; 4498 BP.

XX AAD37258;

XX AC
XX AC
DT 21-AUG-2002 (first entry)

XX Adeno-associated virus vector plasmid, AAV-MCK-3531.

XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.

XX Chimeric - Homo sapiens.
OS Chimeric - Unidentified.

XX WO200183695-A2.

XX 08-NOV-2001.

XX 27-APR-2001; 2001WO-US13677.

QY 900 CTGAGAAATACCTCAGTCAGCCAGAGACGGAAGCATGGCTGGATACTTTGGCCGGTGT 959
Db 1900 CTGAGAAATACCTCAGTCAGCCAGAGACGGAAGCATGGCTGGATACTTTGGCCGGTGT 1959
QY 960 TGGGATAATTTAGTCCAAAACCTTGAAGAAGTACAGCAC 999
Db 1960 TGGGATAATTTAGTCCAAAACCTTGAAGAAGTACAGCACAGACTCAATAGATTACTGCAA 2019
QY 1000 ----- 999
Db 2020 CAGTTCCCTCCGACCTGGAAGAAAGTTCTTGCTGGCTTTACAGAAGCTGAAACAACCTGCC 2079
QY 1000 ----- 999
Db 2080 AATGTCTACAGGATGCTACCCGTAAGGAAGGCTCTAGAGACTCCAAAGGAGTAAAA 2139
QY 1000 ----- 999
Db 2140 GAGCTGATGAAAACAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGATGTTTAT 2199
QY 1000 ----- 999
Db 2200 CACAACCTGGATGAAGACAGCAAAAATCTTGAGATCCCTGGAAGGTTCCGATGATCA 2259
QY 1000 ----- 999
Db 2260 GTCTCTTACAAAAGACGTTTGGATAACATGAACCTTCAAGTGGAGTGAACTTCGGAAGAAAG 2319
QY 1000 ----- 999
Db 2320 TCTCTCAACATTAGGTCCCATTTGGAGCCAGTCTTGACAGTGAAGCGGTCTGCACCTT 2379
QY 1000 ----- 999
Db 2380 TCTCTCAGGAACCTCTGCTGTGGCTTACAGCTGAAAAGATGATGAATTAAGCCGAGCA 2439
QY 1000 ----- 999
Db 2440 CCTATTGGAGGCGACTTTCAGAGCTTACAGAGCAAGATGTATACATAGGCGCTTCAAG 2499
QY 1000 ----- 999
Db 2500 AGGGAATTTGAABAACTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGATATTT 2559
QY 1000 ----- 999
Db 2560 CTGACAGAGCAGCCTTTTGAAGGACTAGAGAAAACCTACAGGAGCCAGAGCTGCCT 2619
QY 1000 ----- 999
Db 2620 CTTGAGGAGAGAGCCAGAAATGTCACCTGGCTTCTACGAAAGAGGCTGAGAGGTCAAT 2679
QY 1000 -----AGACC 1004
Db 2680 ACTGAGTGGGAAAAATTGAACCTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACC 2739
QY 1005 CTTGAAGACTCCAGGAACCTTCAAGAGGCCACGATGAGCTGGACCTCAGCTGGCCAA 1064
Db 2740 CTTGAAGACTCCAGGAACCTTCAAGAGGCCACGATGAGCTGGACCTCAAGCTGGCCAA 2799
QY 1065 GCTGAGGTGATCAAGGATCCTGGCAGCCGCTGGGGATCTCCTCATTTGACTCTCTCCAA 1124
Db 2800 GCTGAGGTGATCAAGGATCCTGGCAGCCGCTGGGGATCTCCTCATTTGACTCTCTCCAA 2859
QY 1125 GATCACTCGAGAAAGTCAAGGACCTTCAAGGAGAAATTTGGCCTCTGAAAGAGAACGTG 1184
Db 2860 GATCACTCGAGAAAGTCAAGGACCTTCAAGGAGAAATTTGGCCTCTGAAAGAGAACGTG 2919
QY 1185 AGCCACGTCAATGACCTTGTCTCCAGCTTACCACTTTGGGATTCAGCTCTCACCGTAT 1244
Db 2920 AGCCACGTCAATGACCTTGTCTCCAGCTTACCACTTTGGGATTCAGCTCTCACCGTAT 2979

QY 1245 AACCTCAGACACTCTGGAACACCTGAACACAGATGAAGAGTCTTCTCAGGTGCCGTGAG 1304
Db 2980 AACCTCAGACACTCTGGAAGACCTGAACACAGATGAAGAGTCTTCTCAGGTGCCGTGAG 3039
QY 1305 GACCGAGTCAGGAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGACATTT 1364
Db 3040 GACCGAGTCAGGAGCTGCATGAAGCCCAAGGAGCTTTGGTCCAGCATCTCAGACATTT 3099
QY 1365 CTTTCCACGCTCTGTCAGGCTCCCTGGGAGAGAGGACATCTGCCAAAACAAAGTGCCCTAC 1424
Db 3100 CTTTCCACGCTCTGTCAGGCTCCCTGGGAGAGAGGACATCTGCCAAAACAAAGTGCCCTAC 3159
QY 1425 TATATCAACACAGAGACTCAAAACAATCTGCTGGGACCATCCCAAATATGACAGACTCTAC 1484
Db 3160 TATATCAACACAGAGACTCAAAACAATCTGCTGGGACCATCCCAAATATGACAGACTCTAC 3219
QY 1485 CAGTCTTTAGCTGACCTGAATATGTCAGATTCAGCTTATAGGACTGCCATGAACTC 1544
Db 3220 CAGTCTTTAGCTGACCTGAATATGTCAGATTCAGCTTATAGGACTGCCATGAACTC 3279
QY 1545 CGAAGACTGCAGAGGCCCTTTGCTTTGGATCTCTTGAGCCTGTCAGCTGCATGTATGCC 1604
Db 3280 CGAAGACTGCAGAGGCCCTTTGCTTTGGATCTCTTGAGCCTGTCAGCTGCATGTATGCC 3339
QY 1605 TTGACACAGACACACTCAAGCAAAATGACCCCATGGATATCTCTCAGATTTATTAAT 1664
Db 3340 TTGACACAGACACACTCAAGCAAAATGACCCCATGGATATCTCTCAGATTTATTAAT 3399
QY 1665 TGTTTGACACACTATTTATGACCGCCTGGAGCAAGAGCAACAATTTGCTCAACGCTCCT 1724
Db 3400 TGTTTGACACACTATTTATGACCGCCTGGAGCAAGAGCAACAATTTGCTCAACGCTCCT 3459
QY 1725 CTCTGCTGATATGTCGTAACCTGGCTGCTGATGTTTATGATACGGGACGACAGGG 1784
Db 3460 CTCTGCTGATATGTCGTAACCTGGCTGCTGATGTTTATGATACGGGACGACAGGG 3519
QY 1785 AGGATCCGCTGCTGCTTTTAAACTGGCATCATTTCCCTGCTGTAAGACACATTTGGAA 1844
Db 3520 AGGATCCGCTGCTGCTTTTAAACTGGCATCATTTCCCTGCTGTAAGACACATTTGGAA 3579
QY 1845 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGCC 1904
Db 3580 GACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGTGACCAAGCC 3639
QY 1905 AGGCTGGCCCTCCTCTGATGATTTCTATCCAAATTCAGAGCAGTTGGGTGAAGTTGCA 1964
Db 3640 AGGCTGGCCCTCCTCTGATGATTTCTATCCAAATTCAGAGCAGTTGGGTGAAGTTGCA 3699
QY 1965 TCCTTTGGGGGACAGTAACATTTGAGCCCAAGTGTCCGGA 2001
Db 3700 TCCTTTGGGGGACAGTAACATTTGAGCCCAAGTGTCCGGA 3736

Search completed: February 1, 2004, 12:05:44

Job time : 515.292 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1002.6	50.1	5952	4	US-09-684-875A-1	Sequence 1, Appli
2	991.6	49.6	13977	4	US-09-481-970B-60	Sequence 60, Appl
C 3	869.8	43.5	19307	3	US-08-836-032A-10	Sequence 10, Appl
C 4	869.8	43.5	19307	3	US-09-427-048A-10	Sequence 7, Appl
5	490.6	24.5	6045	4	US-09-091-501B-7	Sequence 7, Appl
6	490.6	24.5	10320	4	US-09-091-501B-9	Sequence 9, Appl
7	79.4	4.0	200	4	US-09-091-501B-5	Sequence 5, Appl
8	78.6	3.9	200	4	US-09-091-501B-4	Sequence 4, Appl
9	78.6	3.9	200	4	US-09-091-501B-6	Sequence 6, Appl
C 10	74.6	3.7	7218	1	US-08-232-463-14	Sequence 14, Appl
11	44.2	2.2	2574	4	US-09-668-313A-10	Sequence 10, Appl
12	42.8	2.1	1179	4	US-09-107-532A-1186	Sequence 1186, Ap
13	42.8	2.1	1690	4	US-09-620-312D-69	Sequence 69, Appl
14	42.8	2.1	7812	3	US-09-368-590-1	Sequence 1, Appl
C 15	42.4	2.1	2704	3	US-08-857-076-44	Sequence 44, Appl
C 16	42.4	2.1	3499	3	US-08-857-076-43	Sequence 43, Appl
17	39.4	2.0	2223	1	US-08-257-073-4	Sequence 4, Appl
18	38.8	1.9	289	3	US-09-001-005-17	Sequence 17, Appl
19	38.8	1.9	289	3	US-09-244-796-17	Sequence 17, Appl
20	38.8	1.9	1821	4	US-08-477-831C-1	Sequence 1, Appl
21	38.8	1.9	1885	4	US-08-477-831C-9	Sequence 9, Appl
22	38.8	1.9	1896	4	US-08-477-831C-10	Sequence 10, Appl
23	38.8	1.9	1961	4	US-08-477-831C-8	Sequence 8, Appl
24	38.8	1.9	2968	4	US-08-477-831C-13	Sequence 13, Appl
25	38.8	1.9	3044	4	US-08-477-831C-12	Sequence 12, Appl
26	38.4	1.9	7672	4	US-09-220-132-24	Sequence 24, Appl
27	38.2	1.9	428	4	US-09-668-313A-3	Sequence 3, Appl

Db 3919 CTCTCACGTTATACCTCAGCACTCTGGAAGACTGAACACACAGATGGAAGCTTCTGCAG 3978
Qy 1293 GTGCCCTCGAGGACCGAGTCAGGACGTCATGAAGCCCAAGGACCTTGGTCCAGCA 1352
Db 3979 GTGCCCTCGAGGACCGAGTCAGGACGTCATGAAGCCCAAGGACCTTGGTCCAGCA 4038
Qy 1353 TCTCAGCACTTCTTCTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAAC 1412
Db 4039 TCTCAGCACTTCTTCTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGCCATCTCGCCAAAC 4098
Qy 1413 AAAGTGCCCTACTATATCAACACGAGACTCAACAACTCTGCTGGAGACCATCCCAAATG 1472
Db 4099 AAAGTGCCCTACTATATCAACACGAGACTCAACAACTCTGCTGGAGACCATCCCAAATG 4158
Qy 1473 ACAGAGCTCTACAGTCTCTTAGCTGACCTCAATAATGTCAAGTCTCAGCTTATAGACT 1532
Db 4159 ACAGAGCTCTACAGTCTCTTAGCTGACCTCAATAATGTCAAGTCTCAGCTTATAGACT 4218
Qy 1533 GCCATGAACCTCCGAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAAGT 1592
Db 4219 GCCATGAACCTCCGAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGAGCCCTGTCAAGT 4278
Qy 1593 GCATGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACAGCCCATGGATATCCTG 1652
Db 4279 GCATGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACAGCCCATGGATATCCTG 4338
Qy 1653 CAGATTATTAATTTGTTGACCACTATTTATGACCGCTCGAGCAAGAGCAACAAATTTG 1712
Db 4339 CAGATTATTAATTTGTTGACCACTATTTATGACCGCTCGAGCAAGAGCAACAAATTTG 4398
Qy 1713 GTCAAGTCTCTCTCGGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTATGATAG 1772
Db 4399 GTCAAGTCTCTCTCGGTGGATATGTGTCTGAACCTGGCTGCTGAATGTTATGATAG 4458
Qy 1773 GGAGCAACAGGGAGATCCGCTGCTCTCTTTTAAACTGGCATCTTCCCTGTCTAAA 1832
Db 4459 GGAGCAACAGGGAGATCCGCTGCTCTCTTTTAAACTGGCATCTTCCCTGTCTAAA 4518
Qy 1833 GCACATTTGGAAGCAAGTACAGATACCTTTCAAGCAAGTGCGAAGTTCAACAGATTT 1892
Db 4519 GCACATTTGGAAGCAAGTACAGATACCTTTCAAGCAAGTGCGAAGTTCAACAGATTT 4578
Qy 1993 TGTACACAGCGCAGCTGGGCTCTCTCTCATGATTTCTATCAAAATCCAAACAGAGTTG 1952
Db 4579 TGTACACAGCGCAGCTGGGCTCTCTCTCATGATTTCTATCAAAATCCAAACAGAGTTG 4638
Qy 1953 GGTGAAGTTGCATCTTTGGGGCAGTAACATTTGAGCCAAAGTGTCCGGA 2001
Db 4639 GGTGAAGTTGCATCTTTGGGGCAGTAACATTTGAGCCAAAGTGTCCGGA 4687

RESULT 2

US-09-484-970B-60
; Sequence 60, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 60
; LENGTH: 13977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 229357.11CBI
; NAME/KEY: unsure

; LOCATION: 11721-11761, 12294, 13969
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-60

Query Match 49.6%; Score 991.6; DB 4; Length 13977;
Best Local Similarity 99.5%; Pred. No. 2e-297;
Matches 1005; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 993 ACAGCACAGACCTTTGAAGACTCCAGGAATCTCAAGAGGCCACGGATGAGCTGGACCTC 1052
Db 8993 ATAGATGAGACCTTTGAAGACTCCAGGAATCTCAAGAGGCCACGGATGAGCTGGACCTC 9052
Qy 1053 AAGCTGGCCAAAGTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGCATCTCCTCAT 1112
Db 9053 AAGCTGGCCAAAGTGAGGTGATCAAGGGATCTCTGGCAGCCCGTGGCGCATCTCCTCAT 9112
Qy 1113 GACTCTCTCAAGATCACTTCGAGAAAGTCAAGGCACTTCGAGAGAAATTTGCGCTCTG 1172
Db 9113 GACTCTCTCAAGATCACTTCGAGAAAGTCAAGGCACTTCGAGAGAAATTTGCGCTCTG 9172
Qy 1173 AAAGAGAACGTGAGCCACGTCAATGACCTTCTCGCCAGCTTACCACTTTGGGCAATTCAG 1232
Db 9173 AAAGAGAACGTGAGCCACGTCAATGACCTTCTCGCCAGCTTACCACTTTGGGCAATTCAG 9232
Qy 1233 CTCTCAGGTATACCTTCAGCACTCTCGAAGACCTGAAACACCATGGAAGCTTCTCGAG 1292
Db 9233 CTCTCAGGTATACCTTCAGCACTCTCGAAGACCTGAAACACCATGGAAGCTTCTCGAG 9292
Qy 1293 GTGGCCGTGAGGACCGAGTCAGGACGCTCATGAAGCCCAAGGACCTTTGGTCCAGCA 1352
Db 9293 GTGGCCGTGAGGACCGAGTCAGGACGCTCATGAAGCCCAAGGACCTTTGGTCCAGCA 9352
Qy 1353 TCTCAGCACTTCTTCTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGCACTTCGCCAAAC 1412
Db 9353 TCTCAGCACTTCTTCTTCCAGCTCTGTCCAGGCTCCCTGGGAGAGAGCACTTCGCCAAAC 9412
Qy 1413 AAAGTGCCCTACTATATCAACACGAGACTCAACAACTTGTCTGGGACCATCCCAAATG 1472
Db 9413 AAAGTGCCCTACTATATCAACACGAGACTCAACAACTTGTCTGGGACCATCCCAAATG 9472
Qy 1473 ACAGAGCTCTACAGTCTTTAGCTGACCTGAATATGTGAGATCTCAGCTTATAGACT 1532
Db 9473 ACAGAGCTCTACAGTCTTTAGCTGACCTGAATATGTGAGATCTCAGCTTATAGACT 9532
Qy 1533 GCCATGAACCTCCGAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGAGCTGTCAAGT 1592
Db 9533 GCCATGAACCTCCGAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGAGCTGTCAAGT 9592
Qy 1593 GCATGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACAGCCCATGGATATCCTG 1652
Db 9593 GCATGTGATGCTTGGACGAGCAACCTCAAGCAAAATGACAGCCCATGGATATCCTG 9652
Qy 1653 CAGATTATTAATTTGTTGACCACTATTTATGACCGCTCGAGCAAGAGCAACAAATTTG 1712
Db 9653 CAGATTATTAATTTGTTGACCACTATTTATGACCGCTCGAGCAAGAGCAACAAATTTG 9712
Qy 1713 GTCACAGTCCCTCTCTGCTGGATATGTCTGAACCTGGCTGCTGAATGTTATGATAG 1772
Db 9713 GTCACAGTCCCTCTCTGCTGGATATGTCTGAACCTGGCTGCTGAATGTTATGATAG 9772
Qy 1773 GGAGCAACAGGGAGATCCGCTGCTTCTTTTAAACTGCGCATCTTCCCTGTGTAAA 1832
Db 9773 GGAGCAACAGGGAGATCCGCTGCTTCTTTTAAACTGCGCATCTTCCCTGTGTAAA 9832
Qy 1833 GCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGCGAAGTTCAACAGGATTT 1892
Db 9833 GCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGCGAAGTTCAACAGGATTT 9892
Qy 1893 TGTGACACGCGCAGCT--GGGCTCTCTTCTGCAATTTCTATCAAAATCCAAACAGATTT 1951
Db 9893 TGTGACACGCGCAGCTGGGCTCTCTCTGCAATTTCTATCAAAATCCAAACAGATTT 9952
Qy 1952 GGTGAAGTTGCATCTTTGGGGCAGTAACATTTGAGCCAAAGTGTCCGGA 2001

Db 9953 GGGTGAAGTTGCATCCTTTGGGGGCAGTAACATTGAGCCAAAGTGTCGGGA 10002

RESULT 3

US-08-836-022A-10/c
Sequence 10, Application US/08836022A
Patent No. 6001557
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
APPLICANT: Allison, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Chen, Shu-Jen
APPLICANT: Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
City: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,022A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/331,381
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNVPN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-836-022A-10

Query Match	43.5%	Score 869.8	DB 3	Length 19307
Best Local Similarity	91.4%	Pred. No. 2e-259		
Matches 922	Conservative 0	Mismatches 87	Indels 0	Gaps 0
QY	993	ACACGACAGACCCCTTGAAGACATCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACCTC	1052	
Db	5701	ATAGATGAAGCTCTTGAAGAGACTCCAGGAACTTCAAGGAGTGTCGGATGAATCGGACCTC	5642	
QY	1053	AAGCTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGGGATCTCCTCATTT	1112	
Db	5641	AAGTTGGCCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCAGTGGGGGATCTCCTCATTT	5582	
QY	1113	GACTCTCTCCAAAGATCACTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATTTGGCCCTCTG	1172	
Db	5581	GACTCTCTGCAAGATCACTCTGAAAAGTCAAGGCACCTTCGGGGAGAAATTTGCACCTCTTT	5522	
QY	1173	AAAGAGACGTGAGCCACGTCATGACCTTGCTCGCCAGCTTACACATTTGGGGCATTCAG	1232	
Db	5521	AAAGAGAAATGTCAATCTGTCAATGACCTTTGCACATCAGCTGACCCACACTGGGCATTCAG	5462	
QY	1233	CTCTCACGGTATACCTTCAGCACTCTGGGAAGACCTGAAACACGATGGAAGCTTTCTGCAG	1292	

5461	DB	CTCTCACCTTATAA	CCTCAGCAGCTTTGGAGAGTCTGAAATACACAGATGGAGGCTTCTACAG	5404
1293	QY	GTGGCGTTCAGGAG	CCGAGTACAGGCAGCTGCATGAAGCCCAACAGGCACTTTGGTCCAGCA	1352
5401	DB	GTGGCTGTGGAGAG	CCGTTGCACAGCTGCATGAAGCCCAACAGGCACTTTGGTCCCTGCA	5342
1353	QY	TCTCAGCACTTTCTTT	CCACGTCGTCCAGGTCCTCGGAGAGAGCCATCTCGCCAAAC	1412
5341	DB	TCCAGCACTTCTTT	CCACTTCAGTTCCAGGTCCTCGGAGAGAGCCATCTCACCNAAC	5282
1413	QY	AAAGTGCCCTACTAT	ATCAACACAGAGCTCAAAACAACATTGCTGGAGACCATCCCAAAATG	1472
5281	DB	AAAGTGCCCTACTAT	ATCAACACAGAGACCCAAACCACTTGTGGAGACCAACCCCAAAATG	5222
1473	QY	ACAGAGCTCTACAGCT	TTTACTGACCTGAATATATGTACAGATTCTCAGCTTTATAGGACT	1532
5221	DB	ACAGAGCTCTACAGCT	TTTACTGACCTGAATATATGTACAGTTCTCCGCGTATAGGACT	5162
1533	QY	GCATGAAACTCCGAGA	CTGCAGAGGCGCTTTGTGTGATCTCTTTGAGCCTGTGAGCT	1592
5161	DB	GCATGAAGCTCAGAAG	GGCTCCAGAAGGCGCTTTGTGTGATCTCTTTGAGCCTGTGAGCT	5102
1593	QY	GCATGTGATGCTTGGAC	CAGCACAACTCAAGCAAAATGACAGGCCATGGATATCCTG	1652
5101	DB	GCATGTGATGCTTGGAC	CAGCACAACTCAAGCAAAATGACAGGCCATGGATATCCTG	5042
1653	QY	CAGATTATTAATGT	TTTGACCCTATTTATGACCGCTCGAGCAAGAGACACAACATTTG	1712
5041	DB	CAGATAATTAATGT	TTTGACTACATTTATGATCGTCTGGAGCAAGAGCACACAATCTG	4982
1713	QY	GTCAACGTCCTCTCT	CGTGATATGTGTGAACTGGCTGCTGAATGTTATGATACG	1772
4981	DB	GTCAATGTCCCTCTCT	GTGATATGTGTCTCAACTGGCTTCTCAATGTTTATGATACG	4922
1773	QY	GGACGAAACAGGAGGAT	CCGTGCTCTTTTAAACTGGGATCATTTCCCTGTCTGATAA	1832
4921	DB	GGACGAAACAGGAGGAT	CCGTGCTCTTTTAAACTGGGATCATTTCTCTGUGTATAA	4862
1833	QY	GCACATTTGGAGAGA	CAAGTACATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTT	1892
4861	DB	GCACATTTGGAGAGA	CAAGTACATACCTTTTCAAGCAAGTGGCAAGTTCAACTGCTGCTT	4802
1893	QY	TGTGACCAACGAGCT	GGGCGCTCCTCTGATGATTTCTATCCAATTTCCAAGACAGTTG	1952
4801	DB	TGTGACCAACGAT	AGCTGGGCTCTTTCTGCAATTTCTATTTCAATTTCCAAGACAGTTG	4742
1953	QY	GGTGAAGTTGCTAC	CTCTTTGGGGCAGTAACTTTGAGCCCAAGTGTCCGGA	2001
4741	DB	GGTGAAGTTGCTCT	TTTGGGGCAGTAACTTTGAGCCCAAGTGTGAGGA	4693

RESIST. 4

US-09-427-048A-10/c
Sequence 10, Application US/09427048A
Patent No. 6203975
GENERAL INFORMATION:
APPLICANT: Trustees of the University of Pennsylvania
Wilson, James M.
Fisher, Krishna J.
Chen, Shu-Jen
Weitzman, Matthew
TITLE OF INVENTION: Improved Adenovirus Virus and
Methods of Use Thereof
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P O Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/427,048A
FILING DATE: 21-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,022
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GVPN.008PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19307 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-427-048A-10

Query Match 43.5%; Score 869.8; DB 3; Length 19307;
Best Local Similarity 91.4%; Pred. No. 2e-259;
Matches 922; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 993 ACAGCAGACCTTGAAGACCTCCAGGACTTCAAGAGCCACGGATGAGTGGACCTC 1052
DB 5701 ATAGATGAAGCTCTTGAAGACCTCCAGGACTTCAAGAGCTCCGATGAATGGACCTC 5642
QY 1053 AAGTCGCCAAGCTGAGGTGATCAAGGATCTCGCAGCCCGTGGCGATCTCCCTATT 1112
DB 5641 AAGTTCGCCAAGCTGAGGTGATCAAGGATCTCGCAGCCCGTGGCGATCTCCCTATT 5582
QY 1113 GACTCTTCAAGATCACTTCGAGAAAGTCAAGGACTTCGAGGAGAAATGGGCTCTG 1172
DB 5581 GACTCTTCAAGATCACTTCGAGAAAGTCAAGGACTTCGAGGAGAAATGGGCTCTT 5522
QY 1173 AAAGAGAACGTGAGCCACGTCAATGACCTTGCCTCGCAGCTTACCACTTTGGCATTTCAG 1232
DB 5521 AAAGAGATGTCAATCTGTGTAATGACCTTGCATCAGTGCACACACTGGCATTTCAG 5462
QY 1233 CTCTCACCGTATAACCTCAGACTCTGGAAGACTGAAACACCAAGATGGAGCTTCTGCAG 1292
DB 5461 CTCTCACCTTATAACCTCAGACTCTTGAAGATCTGAATACCAAGATGGAGCTTCTPACAG 5402
QY 1293 GTGSCCGTCGAGGACCGAGTCAGCAGCTGCATGAAGCCACAGGACTTTGGTCCAGCA 1352
DB 5401 GTGSCGTGGAGGACCGGTGTCAGACAGCTGCATGAAGCCACAGGACTTTGGTCTCGCA 5342
QY 1353 TCTCAGCACTTTCTTCCAGCTCTGTCCAGGTCCTTGGGAGAGAGCCATCTCGCCAAAC 1412
DB 5341 TCCAGCACTTTCTTCCACTTCAGTTTCAGGTCCTTGGGAGAGAGCCATCTCAACCAAC 5282
QY 1413 AAAGTGCCTACTATATCAACACAGAGACTCAACAACTTGTGGAGCCATCCCAAAATG 1472
DB 5281 AAAGTGCCTACTATATCAACACAGAGACTCAACAACTTGTGGAGCCATCCCAAAATG 5222
QY 1473 ACAGAGCTTACCAAGTCTTTAGTGTGACTGAATATGTCAGATTCTCAGCTTATAGGACT 1532
DB 5221 ACAGAGCTTACCAAGTCTTTAGTGTGACTGAATATGTCAGATTCTCCGCGTATAGACT 5162
QY 1533 GCCATGAATTCGGAAGACTTCGAGAGCCCTTTGTGTGATCTCTTGTAGCCTGTGAGCT 1592
DB 5161 GCCATGAAGTCTCAGAGGCTCCAGAGGCTTTGTGTGATCTCTTGTAGCCTGTGAGCT 5102
QY 1593 GCATGTGATGCTTGGACCAAGCAACTCAAGCAAAATGACCAAGCCCATGGATATCTCTG 1652

DB 5101 GCATGTGATGCTTGGACCAAGCAACTCAAGCAAAATGACCAAGCCCATGGATATCTCTG 5042
QY 1653 CAGATTATTAATTTGTTGACCACTATTATGACCGCTTGGAGCAAGAGCAACAATTTG 1712
DB 5041 CAGATAATTAATTTGTTGACTACATATTAATGATGCTTGGAGCAAGAGCAACAATCTG 4982
QY 1713 GTCAACGTCCTCTCTCGTGGATATGCTGAACTGCTGCTGCTGCTGAATGTTATGATACG 1772
DB 4981 GTCAATGTCCCTCTCTCTGTGGATATGCTCAACTGGCTTCTCAATGTTTATGATACG 4922
QY 1773 GGAGCAACAGGAGGATCCGTCCTCTCTTTTAAACTGSCATCATTTCCCTGTGTAAA 1832
DB 4921 GGAGCAACAGGAGGATCCGTCCTCTCTTTTAAACTGSCATCATTTCTCTGTGTAAA 4862
QY 1833 GCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTT 1892
DB 4861 GCACATTTGGAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACTGGCTTT 4802
QY 1893 TGTACACAGCGCAGGCTGGGCTCTCTTCTGATGATTTCTATCCAAATCCCAAGACAGTTG 1952
DB 4801 TGTACACAGCGTAGGCTGGGCTCTTCTTCTGATGATTTCTATCCAAATCCCAAGACAGTTG 4742
QY 1953 GGTGAAGTTGCTATCTTTGGGGGCGAGTAACTTGAGCAAGTGTCCGGA 2001
DB 4741 GGTGAAGTTGCTATCTTTGGGGGCGAGTAACTTGAGCGGAGTGTTCAGGA 4693
RESULT 5
US-09-091-501B-7
; Sequence 7, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(6037)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
; US-09-091-501B-7
Query Match 24.5%; Score 490.6; DB 4; Length 6045;
Best Local Similarity 67.6%; Pred. No. 9.7e-142;
Matches 588; Conservative 0; Mismatches 329; Indels 0; Gaps 0;
QY 985 AAAAGAGTACAGCAGACCCCTTGAAGACTTCCAGGAACCTTCAAGAGCCACCGATGACC 1044
DB 3783 AAAAGCAAGTGCACAGGCAATTTGAGAAACTCAGAGACCTCGAGGGAGCTATGATGACC 3842
QY 1045 TGGACCTCAAGCTGGCCAGCTGAGTGTATCAAGGATCTTGCAGCCCGTGGCGATC 1104

3843 TGGACGCTGACATGAAGGAGGAGAGTCCGTGCGGAATGGCTGGAAGCCCGTGGAGACT 3902
1105 TCCTCATTGACTCTCTCAAGATCACTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTG 1164
3903 TACTCATTGACTCGTGGAGGATCATTTGAAATATCATGGCATTTAGAGAAGAAATTG 3962
1165 CCGCTCTGAAGAGAACGTGAGCCAGCTCAATGACCTTTCGCGAGCTTACCACTTTGG 1224
3963 CACCAATCAACTTTAAAGTTAAACGGTGAATGATTTATCCAGTCAGCTGTCTCCACTTG 4022
1225 GCATTGAGCTCTCCCGGTATACCTCAGCACTCTGGAAGACCTGAACACACAGATGGAAGC 1284
4023 ACTGATCCCTCTCTTAAAGATGTCTCCGAGCTAGATGACCTTAATGCGAGAAAC 4082
1285 TTCTCAGGTGCGGCTGAGGACCGAGTCAGGACGCTGCATGAAGCCACACAGGACTTTG 1344
4083 TTTTACAGGTTTCTGTGATGATGCGCTTAAACAGCTTCAGGAAGCCACACAGATTTTG 4142
1345 GTCCAGATCTCAGCACTTTCTTCCAGCTGTGCCAGGTCCCTGGGAGAGCCATCT 1404
4143 GACCATCTCTCAGCACTTTCTCTACGTGCTCCAGCTCCGCTGGCAAGATCCATTT 4202
1405 CCGCAAAACAAAGTGCCTTACTATATCAACACGAGACTCAAAACAACTTGTGGAGCCATC 1464
4203 CACATAATAAGTGCCTTATACATCAACCATCAACACAGACCACTTGTGGAGCCATC 4262
1465 CCAAAATGACAGAGCTTACAGCTCTTTCAGCTCTTTCAGCTTTCAGATTTCTCAGCTT 1524
4263 CTAATAATGACGAACTCTTCAATCCCTTGTGACCTGAAATATGATGCTTTCGCT 4322
1525 ATAGAGCTGCCATGAACCTCGAGAGCTGAGAGGCGCTTGTCTGGATCTTTCGAGCC 1584
4323 ACCGTACAGCAATCAAAATCCGAGACTCAAAAGGCACTATGTTGGATCTCTTAGAGT 4382
1585 TGTGAGCTGCTGATGCTTGGACGACGACCAACCTCAAGCAAAATGACAGCCCATGG 1644
4383 TGAGTACAAACAAATGAATTTTCAACAGCACTAGTTGNACCAAAATGACAGCTCTCTCA 4442
1645 ATATCTCGAGATTAATTAATTTGTTGACCACTATTATGACCGCTGGAGAGAGACACA 1704
4443 GTGTCCAGATGTCTCACTGCTGACCACTATGATGAGTGTGATGCTCAATGCTCAATGCT 4502
1705 ACAATTTGGTCAAGCTGCTCTCGGTGGATGATGCTGAACTGCTGAACTGCTGAACTGTTT 1764
4503 AGGACCTGCTCAAGTTCACCTCTGTGTGATGATGCTCAATGCTGCTCAATGCTCAATGCT 4562
1765 ATGATACGGGACGACGAGGAGGATCCGTGCTCTTTTAAACTGGCATCATTTCC 1824
4563 ATGACACGGGTGGAATTTAGAGTGCAGAGTCTGAAGATTGGATTAAATGCTC 4622
1825 TGTGTAAGACATTTGGAACAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAA 1884
4623 TCTCAAGAGTCTCTTGAAGAAAATACAGATATCTCTTTAAGGAAGTTGCGGGGCCGA 4682
1885 CAGGATTTTGTACACGCGAGGCTGGGCTCTCTTTCGATGATTTCTATCAAAATCCAA 1944
4683 CAGAAATGTGTGACCAAGGCGAGTGGGCTGTACTTCAATGATGATGATGATGATGATGATG 4742
1945 GACAGTTGGGTGAAGTGTGATCTCTTTGGGGGAGTAACTTGGAGCCAAAGTGTCCGGA 2001
4743 GCGAGCTAGGTGAAGTAGAGCTTTTGGAGGCAATATTTAGGCTAGTGTGTTCCGA 4799

RESULT 6
US-09-091-501B-9
; Sequence 9, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091.501B

; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 10320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(10312)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (724)..(758)
; OTHER INFORMATION: Precise residue is left open
; US-09-091-501B-9
Query Match 24.5%; Score 490.6; DB 4; Length 10320;
Best Local Similarity 67.6%; Pred. No. 1.4e-141;
Matches 688; Conservative 0; Mismatches 329; Indels 0; Gaps 0;
QY 985 AAAAGAGTACAGCAGACAGCCCTTGAAGACTTCAGAGACTTCAGAGCCAGGATGAGC 1044
Db 8058 AAAAGCAAGTGGACAGGCAATTTGGAGAACTCAGAGACTTCAGAGGAGCTATGATGACC 8117
QY 1045 TGGAGCTCAAGCTGCGCCAGCTGAGGTGATCAAGGGATCTCGGAGCCCGTGGCGATC 1104
Db 8118 TGGAGCGTGCATGAAGGAGGAGGAGTCCGTGCGGAATGGCTGGAAGCCCGTGGAGACT 8177
QY 1105 TCCTCATTGACTCTCTCAAGATCACTTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTG 1164
Db 8178 TACTCATTGACTCGTGCAGGATCACAATTGAAATAATCATGGCATTTAGAGAAGAAATTG 8237
QY 1165 CCGCTCTGAAGAGAACGTGAGCCAGCTCAATGACCTTGTGCGCAGCTTACCAGCTTTGG 1224
Db 8238 CACCAATCAACTTTAAAGTTAAACCGGTGAATGATTTATCCAGTCAGCTGTCTCCACTTG 8297
QY 1225 GCATTGAGCTCTCACCGGTATAACCTCAGCACTCTGGAAGACTGAAACACAGATGGAAGC 1284
Db 8298 ACTGATCCCTCTCTAAAGATGTCTGCCAGCTAGATGACCTTAATATGCGATGGAAC 8357
QY 1285 TTCTCAGGTGCGGCTGAGGACCGAGTCAAGGAGTGCATGAAGCCACAGGAGCTTTG 1344
Db 8358 TTTTACAGGTTTCTGTGATGATGCGCTTAAACAGCTTCAGGAAGCCACACAGATTTTG 8417
QY 1345 GTCCAGATCTCAGCACTTTCTTCCAGCTGTGCGAGGTCCTGGGAGAGGCCATCT 1404
Db 8418 GACCATCTCTCAGCACTTTCTCTCTAGTCAGTCCAGCTCCGCTGGCAAGATCCATTT 8477
QY 1405 CCGCAAAACAAAGTGCCTTACTATATCAACACGAGACTCAAAACAACTTGTCTGGAGCCATC 1464
Db 8478 CACATAATAAGTGCCTTATACATCAACCATCAACACAGACCACTGTTGGAGCCATC 8537
QY 1465 CCAAAATGACAGAGCTTACAGCTTTTAGTGCATGATTAATGTCAGATTTCTCAGCTT 1524
Db 8538 CTAATAATGACCGAACTCTTCAATCCCTTGTCTGACCTGAAATGATGATGATGATGATG 8597
QY 1525 ATAGAGCTGCCATGAACCTCCGAAAGACTTCGAGAGGCGCTTGTCTGGATCTCTTGGAGCC 1584
Db 8598 ACCGTACAGCAATCAAAATCCGAGACTTACAAAAGCACTATGTTGGATCTCTTAGAGT 8657
QY 1585 TGTGAGCTGATGATGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1644

TYPE: DNA
ORGANISM: Homo sapiens
US-09-091-501B-6

Query Match 3.9%; Score 78.6; DB 4; Length 200;
Best Local Similarity 62.4%; Pred. No. 1.6e-14;
Matches 123; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 452 TGAATCGAAGACCTTAAACCCCAAGTACAAACATAGGTGCTTCAAGAAAGATCTAGA 511
Db 4 TGATGTGAATCTCTACAAAAGCTGTAGAAAGACATAAAAGTTTGCAAGTGTATCTGA 63
QY 512 ACAAGAAAGTCAGGTCATCTCTCACTCACATGCTGTGTAGTGTATGATCTAG 571
Db 64 GGCTGAACAGGTGAAGTAAATTCACCTAACTCACATGCTGTGCTATTGTTGATGAAGACAG 123
QY 572 TGGAGATCACCAACTGCTGCTGTTGGAAAGAACAACTTAAAGTATTTGGAGATCGATGGC 631
Db 124 TGGTGAGAGCGCTACAGCTATCTTAGAAGACCAAGTACAGAACTTGGTGAGCGCTGAC 183
QY 632 AAACATCTGTAGATGGA 648
Db 184 AGCAGTATGCCGTTGGA 200

RESULT 10
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F15
US-08-232-463-14

Query Match 3.7%; Score 74.6; DB 1; Length 7218;
Best Local Similarity 5.6%; Pred. No. 3.4e-12;
Matches 23; Conservative 236; Mismatches 150; Indels 0; Gaps 0;

QY 113 TTCTTAATGATGTGGAAGTGGTGAAGACCACTTTTCATCTACTCATGAGGGGTACATGATGG 172
Db 1474 TATCTATGCAAGTAGTTAAAGAGATAGAGAAATTTGGTACRRRRRRRRRRRRRRRR 1415
QY 173 ATTGACACCCCATCAGGCCGGGTGGTAAATATTCTCAATTTGGGAAGTAAGCTGATTG 232
Db 1414 RRR 1355
QY 233 GAACAGGAAATTTATCAGAAGATGAAGAACTGAAGTACAAAGACAGATGAATCTCCTAA 292
Db 1354 RRR 1295
QY 293 ATTCAAGATGGGAATGCCCTCAGGCTAGCTAGCATCGAAACAAACAAAGCAATTTACATAGAG 352
Db 1294 RRR 1235
QY 353 TTTTAATGGATCTCCAGAAATCGAAATCGAAAGAGTTGAATGCTGCTAACAAACAAAGA 412
Db 1234 RRR 1175
QY 413 AGAAGAGACAAAGAAATGCGAGGAAGACCTCTTGGACCTGATCTTGAAGACCTAAACG 472
Db 1174 RRR 1115
QY 473 CCAAGTACAAACATTAAGGTCTTCAAGAGATCTAGAAACAAAGAACAA 521
Db 1114 RRR 1066

RESULT 11
US-09-668-313A-10
; Sequence 10, Application US/09668313A
; Patent No. 6503756
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF SYNTAXIN 4 INTERACTING PROTEIN EXPRESSION
; FILE REFERENCE: RTS-0127
; CURRENT APPLICATION NUMBER: US/09/668,313A
; CURRENT FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 247
; SEQ ID NO 10
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)...(1891)
US-09-668-313A-10

Query Match 2.2%; Score 44.2; DB 4; Length 2574;
Best Local Similarity 54.7%; Pred. No. 0.0048;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 1389 TGGGAGAGGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACGAGACTCAACA 1448
Db 1733 TGGGAGGAAGCTTACACAGCAGATGGAATCAAGTACTTTCATCAACACGAGTACAGAC 1792
QY 1449 ACTTGTGGGACCATCCCAAAATGACAGAGCTTACCAAGTCTTTAGCTGACCTGAATAAT 1508
Db 1793 AGTCTCTGGATCCACCCCGTGTAGAGCCCTTGACCTGTCTGTGTGAGAGAGAGTAA 1852
QY 1509 GTCAGATTTCTAGCTTATAGAGCTGCCATGAACCTCGAAG 1549
Db 1853 GAGGACTGTCCAGAGAGTAAACAGACCCGAAAGCTGATG 1893

RESULT 12

US-09-107-532A-1186
; Sequence 1186, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GPC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1186:
; LENGTH: 1179 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1179
; SEQUENCE DESCRIPTION: SEQ ID NO: 1186:
US-09-107-532A-1186
Query Match 2.1%; Score 42.8; DB 4; Length 1179;
Best Local Similarity 49.5%; Pred. No. 0.0076;
Matches 110; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 353 TTTTAAATCGATCTCCAGATCAATCAAGAGTTCAAGTGAATGCTGCTAACAACAAACAGA 412
Db 585 TGTTGATCCAGCAGTTGAGCAAAAGATAAAGATTTGATCGATCAGACAAAGAAAATGG 644
QY 413 AGAAGAACACAGGAAATGAGGAGAGACCTCTTTGGACCTGATCTTGAAGACCTAAACG 472
Db 645 AGATACGATCGAGGAATTTGAGAGTGTCTGTTGAGGCGTTCACGCTGGATTAGGAAG 704
QY 473 CCAAGTACAAACATAGTGGTCTCAAGAGATCTAGAACAGACAAAGTCAAGGTCAA 532
Db 705 CTACGTACAATGGGACACAGCTAGATGCCAAATCGCAAGCTGTGTTAGTATCAA 764
QY 533 TTCTCTCACTCACATGGTGGTGTAGTTGATGAATCTAGTGG 574
Db 765 TGCCTTAAAGCGGTAGNATTTGGGTCGGATTCACTTCTGG 806

RESULT 13

US-09-620-312D-69
; Sequence 69, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 69
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1522)
US-09-620-312D-69
Query Match 2.1%; Score 42.8; DB 4; Length 1690;
Best Local Similarity 47.9%; Pred. No. 0.0097;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
QY 1016 CCAGGAATTCAGAGCCACGAGATGAGCTGAGCTCAAGCTGGCCGCAAGTGAAGTGTAT 1075
Db 55 CAAGGAGTTGACCAAGTGGCGCACGACCTGGACGAGCTGGCATGGGTTCAGGAGCG 114
QY 1076 CAAGGGA---TCCTGGCAGCCGCGGCGGATCTCTCATTTGACTCTCTCAAGATCACCT 1132
Db 115 GCTGCCACTGGCCATGACAGACGAGCGAAGCGTTTGCAGCGGTTCAGCGACAT 174
QY 1133 CGAGAAATCAAGGCACCTTCGAGGAGAAATTCGCCCTCTTGAAGAGAACGTTGAGCCACGT 1192
Db 175 CAAAAGAACACAGGCGCTTCGCGCGGAGATCCAGGCGCATGGGCGCGCTTCGAGGAGGT 234
QY 1193 CAATGACCTTCTGCCAGCTTACACTTTGGGCAATTCAGCTCTCCAGGTATAACCTCAG 1252
Db 235 GCTGAGCGCGCGCGCGCTGGCGTCTGCGCAGCCGAGGAGGAGGAGGAGTGGCCCG 294
QY 1253 CACTCTGGAACACCTGAACACACAGATGGAAGCTTCTGCAGGTGGCCCTCGAGGACCGAGT 1312
Db 295 GGGCTTGAGCAGCTGCAGACGCGCTGGCGCGGACTCGGGAGGCTGCCGAGCGACGCA 354
QY 1313 CAGGCGCTGCATGAAGCCACAGG 1338
Db 355 GCAGGTGCTGGACGCGCGCTTCCAGG 380

RESULT 14

```
US-09-368-590-1
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(102)
; NAME/KEY: unsure
; LOCATION: (1021)...(1023)
; NAME/KEY: unsure
; LOCATION: (2266)...(2268)
; US-09-368-590-1

Query Match      2.1%; Score 42.8; DB 3; Length 7812;
Best Local Similarity 47.9%; Pred. No. 0.028;
Matches 156; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 1016 CCAGGAACCTTCAAGAGCCACGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGGTGAT 1075
DB 3639 CAAGGAGTTGCACCGAGTGGCGGCACGACCTGGACGAGCTGCGATGGTTCAGGAGCG 3698
QY 1076 CAAGGGA---TCCTGGGAGCCGCTGGGCGATCTCTCTATTGACTCTCTCCCAAGATCACCT 1132
DB 3699 GCTGCCACTGCCCATGACAGACAGGACGAGCAAGGTTTCAGCGCGTCCAGCGACAT 3758
QY 1133 CGAGAAAGTCAAGGCACCTTGAGGAGAGAAATGGCCCTCTGAAAAGAGAACTGAGCCACGT 1192
DB 3759 CAAAAGAAACACAGGCGCTGCGCGGGAGATCCAGGCGCATGGGCGCGCTGGAGGAGGT 3818
QY 1193 CAATGACCTTCTCGCCAGCTTACCACTTTGGGCAATTCAGCTCTCACCGTATACCTCAG 1252
DB 3819 GCTGGAGCGCGCGCGCGCTGGCGTCTGCTGCGCAGCCCGAGCGAGGCAAGTGGCCG 3878
QY 1253 CACTCTGGAAGACCTGAACACACAGATGGAAGCTTCTGAGTGGCGCTGAGGACCGAGT 1312
DB 3879 GGGCTGGAGCAGCTGCAGAGCGCTGGCGCGGACTCGGGAGGCTGCCGAGCGACGGCA 3938
QY 1313 CAGGAGCTGCATGAAGCCACAGG 1338
DB 3939 GCAGGTGCTGACGCGCGCTCCAGG 3964

Search completed: February 2, 2004, 06:16:03
Job time : 125.697 secs
```

```
FILE REFERENCE: 00786/351001
; CURRENT APPLICATION NUMBER: US/08/857,076C
; CURRENT FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; US-08-857-076-44

Query Match      2.1%; Score 42.4; DB 3; Length 2704;
Best Local Similarity 48.4%; Pred. No. 0.018;
Matches 118; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 196 GTTGGTAATATTCTACAATTGGGAAGTAAAGCTGATTGGAACAGGAAAAATTATCAGAAGAT 255
DB 2116 GCTAGACATTTCTACCGGGGGGAAAAAAGAGAGAAAAATTAAAGAATTACAGAGAAAT 2057
QY 256 GAAGAAACTGAAGTACAAGAGCAGATGAATCTCTTAATTCAGATGGGAATGCCTCAGG 315
DB 2056 CAAGAAATCAAAATTTTTCAAAAATTCAAATTAAGAAAGACAGATTTGGGGAATGTCTTTA 1997
QY 316 GTAGCTAGCATGAAAAACAAGCAATTTACATAGAGTTTTTAATGATCTCCAGATCGA 375
DB 1996 GCAGTTTGTATTCGGAACAATTTACAGCGAAAAAAGATGGGATAGAGTAGCATTTAAA 1937
QY 376 AACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAAAGAAACAAAGGAAAAATGAGG 435
DB 1936 GAGTGAAGAACCTGGATGACGCGGAGAGAGAGAGAAAAAACAAGAAAAAAGAGGT 1877
QY 436 AAGA 439
DB 1876 AATA 1873
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 23:51:42 ; Search time 655.123 Seconds
(without alignments)
11131.886 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000

Perfect score: 2001

Sequence: 1 ggcagtcattgatggagag.....cattgagcgaagtgcgga 2001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	ID	Description
1	2001	100.0	3446	13	US-09-845-416-14
2	1990	99.5	4414	13	US-09-845-416-32
3	1820.8	91.0	5417	13	US-10-149-736-39
4	1593	79.6	3510	13	US-09-845-416-12
5	1593	79.6	4476	13	US-09-845-416-31
6	1263	63.1	3858	13	US-09-845-416-9
7	1263	63.1	4825	13	US-09-845-416-29
8	1263	63.1	4848	13	US-09-845-416-35
9	1263	63.1	5060	13	US-09-845-416-36
10	1254	62.7	5339	13	US-10-149-736-40
11	1251.2	62.5	3531	13	US-09-845-416-10
12	1251.2	62.5	4498	13	US-09-845-416-30
13	1245	62.2	4182	13	US-09-845-416-2
14	1245	62.2	5149	13	US-09-845-416-27
15	1132.2	56.6	5462	13	US-10-149-736-41

Sequence 6, Appl
Sequence 28, Appl
Sequence 34, Appl
Sequence 13, Appl
Sequence 4, Appl
Sequence 42, Appl
Sequence 1, Appl
Sequence 44, Appl
Sequence 47, Appl
Sequence 22, Appl
Sequence 2284, Ap
Sequence 1, Appl
Sequence 108, App
Sequence 15, Appl
Sequence 3, Appl
Sequence 2, Appl
Sequence 7, Appl
Sequence 4, Appl
Sequence 23, Appl
Sequence 3, Appl
Sequence 667, App
Sequence 35, Appl
Sequence 11, Appl
Sequence 8, Appl
Sequence 33, Appl
Sequence 10, Appl
Sequence 9, Appl
Sequence 34, Appl
Sequence 15766, A
Sequence 21956, A

16 1112 55.6 3999 13 US-09-845-416-6
17 1112 55.6 4966 13 US-09-845-416-28
18 1112 55.6 4990 13 US-09-845-416-34
19 1002.6 50.1 1821 13 US-09-845-416-13
20 1002.6 50.1 2169 13 US-09-845-416-4
21 1002.6 50.1 8689 13 US-10-149-736-42
22 1002.6 50.1 11058 13 US-09-845-416-1
23 1002.6 50.1 11443 13 US-10-149-736-44
24 1002.6 50.1 12057 13 US-10-149-736-47
25 1002.6 50.1 13957 10 US-09-782-378A-22
26 1002.6 50.1 13957 10 US-09-880-107-2284
27 1002.6 50.1 13957 13 US-10-149-736-1
28 1002.6 50.1 14082 13 US-10-341-434-108
29 1001 50.0 1434 13 US-09-845-416-15
30 985.8 49.3 1391 13 US-09-845-416-3
31 869.8 43.5 13815 13 US-10-149-736-2
32 665 33.2 1667 13 US-09-845-416-7
33 506.6 25.3 11086 13 US-10-149-736-4
34 490.6 24.5 10302 10 US-09-782-378A-23
35 490.6 24.5 10302 13 US-10-149-736-3
36 481.8 24.1 16531 13 US-10-101-510-667
37 457 22.8 887 13 US-10-149-736-35
38 350 17.5 1340 13 US-09-845-416-11
39 327 16.3 327 13 US-10-149-736-8
40 324 16.2 324 13 US-10-149-736-33
41 322 16.1 333 13 US-10-149-736-10
42 321 16.0 333 13 US-10-149-736-9
43 216 10.8 216 13 US-10-149-736-34
44 180 9.0 476 9 US-09-864-761-15766
45 160.2 8.0 256 9 US-09-864-761-21956

ALIGNMENTS

RESULT 1

US-09-845-416-14
; Sequence 14, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DEL1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; TYPE: DNA
; LENGTH: 3446
; ORGANISM: Homo sapiens
US-09-845-416-14

Query Match 100.0%; Score 2001; DB 13; Length 3446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 60
Db 1000 GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAAACAGCTTTAGAGAA 1059

QY 61 GTATTATCGTGCTCTTCTCTGAGGACACATTCGAGCAGCAGGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGCTCTTCTCTGAGGACACATTCGAGCAGCAGGAGAGATTTCTAAT 1119

QY 121 GATGTGGAAGTGGTGAAGACACAGCTTTTCATCTCATGAGGGGTACATCATGATGATTGACA 180
Db 1120 GATGTGGAAGTGGTGAAGACACAGCTTTTCATCTCATGAGGGGTACATCATGATGATTGACA 1179

QY 181 GCCCATCAGGCGCGGGTGGTGAATATTCTTACATTTGGGAAGTAACCTGATTGGAAACAGGA 240

Db 1180 GCCCATCAGGCCGGTGGTGAATATTTCTCAATTTGGGAAGTAAGCTGATTTGAACAGGA 1239
Qy 241 AAATTATCAGAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTAAATCAAGA 300
Db 1240 AAATTATCAGAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTAAATCAAGA 1299
Qy 301 TGGGAATGCTCAGGTTAGCTAGCTGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 1300 TGGGAATGCTCAGGTTAGCTAGCTGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 1359
Qy 361 GATCTCCAGAAATCAAACTGAAAGAGTTGAATGACTGGCTTAAACAAAACAGAGAAAGAA 420
Db 1360 GATCTCCAGAAATCAAACTGAAAGAGTTGAATGACTGGCTTAAACAAAACAGAGAAAGAA 1419
Qy 421 CAAGGAAATGGAGGAGAGAGCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTAC 480
Db 1420 CAAGGAAATGGAGGAGAGAGCTCTTGGACCTGATCTTGAAGACCTAAACGCCAAGTAC 1479
Qy 481 AACAACTAAGGTGCTTCAAGAGATCTAGAAACAAAGCAAGTCAGGGTCAATTTCTCTCA 540
Db 1480 AACAACTAAGGTGCTTCAAGAGATCTAGAAACAAAGTCAGGGTCAATTTCTCTCA 1539
Qy 541 CTCACTGTTGGTGGTGAATGATGAATCTAGTGAAGATCAAGCACTGCTGCTTTGGAAG 600
Db 1540 CTCACTGTTGGTGGTGAATGATGAATCTAGTGAAGATCAAGCACTGCTGCTTTGGAAG 1599
Qy 601 AACAACTAAGGTGCTTCAAGAGATCTAGAAACAAAGCAAGTCAGGGTCAATTTCTCTCA 660
Db 1600 AACAACTAAGGTGCTTCAAGAGATCTAGAAACAAAGTCAGGGTCAATTTCTCTCA 1659
Qy 661 GGGTCTTTTCAAGACATCTTCTCAAAATGGCAAGCTTTTAAAGCGGATCTAG 840
Db 1660 GGGTCTTTTCAAGACATCTTCTCAAAATGGCAAGCTTTTAAAGCGGATCTAG 1719
Qy 721 TTAGTGCATGGCTTTCAGAAAGAGATGAGTGAACAGATCTCAGCAAGTGGCTTTA 780
Db 1720 TTAGTGCATGGCTTTCAGAAAGAGATGAGTGAACAGATCTCAGCAAGTGGCTTTA 1779
Qy 781 AAGATCAAAATGAATGTTTCAAGTCTTCAAAATGGCGCTTTTAAAGCGGATCTAG 840
Db 1780 AAGATCAAAATGAATGTTTCAAGTCTTCAAAATGGCGCTTTTAAAGCGGATCTAG 1839
Qy 841 AAAAGAAAAGCAATCAATGGGCAAACTGATTTACCTCAAAACAGATCTTCTTTCAACAC 900
Db 1840 AAAAGAAAAGCAATCAATGGGCAAACTGATTTACCTCAAAACAGATCTTCTTTCAACAC 1899
Qy 901 TGAAGATAAGTCAAGTCAACCCAGAGACGGAAGCATGGCTGGATAAGCTTTGCCCGGTGT 960
Db 1900 TGAAGATAAGTCAAGTCAACCCAGAGACGGAAGCATGGCTGGATAAGCTTTGCCCGGTGT 1959
Qy 961 GGGATAATTTAGTCCAAAACCTTGAAAGAGTACAGCACAGACCCCTTGAAGAGCTCCAGG 1020
Db 1960 GGGATAATTTAGTCCAAAACCTTGAAAGAGTACAGCACAGACCCCTTGAAGAGCTCCAGG 2019
Qy 1021 AACTTCAGAGGCCACCGATCAGTGGACCTCAAGCTCGCCAGCTGAGTGAATCAAGG 1080
Db 2020 AACTTCAGAGGCCACCGATCAGTGGACCTCAAGCTCGCCAGCTGAGTGAATCAAGG 2079
Qy 1081 GATCTGCGACCGCTGGGCGATCTCTCAATGACTCTCTCAAGATCACTCGAGAAAG 1140
Db 2080 GATCTGCGACCGCTGGGCGATCTCTCAATGACTCTCTCAAGATCACTCGAGAAAG 2139
Qy 1141 TCAAGGCACTTCAGAGAGAAATGGGCTCTGAAAGAGACGTGAGCCAGTCAATGACC 1200
Db 2140 TCAAGGCACTTCAGAGAGAAATGGGCTCTGAAAGAGACGTGAGCCAGTCAATGACC 2199
Qy 1201 TTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAAGCTCAGCACTCTGG 1260
Db 2200 TTGCTCGCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAAAGCTCAGCACTCTGG 2259
Qy 1261 AAGACTGAACACAGATGGAAGCTTTGCAAGTGGCGCTCGAGGACCGAGTCAGGAGC 1320

Db 2260 AAGACTGAACACACAGATGGAAGCTTCTGCAGGTGGCGCTCGAGGACCGAGTCAGGACG 2319
Qy 1321 TGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTC 1380
Db 2320 TGCATGAAGCCCAACAGGACTTTGGTCCAGCATCTCAGCACTTTCTTTCCAGCTCTGTC 2379
Qy 1381 AGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGA 1440
Db 2380 AGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACAGAGA 2439
Qy 1441 CTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACC 1500
Db 2440 CTCAAAACAACTTGTCTGGGACCATCCCAAAATGACAGAGCTCTACCACTCTTTAGCTGACC 2499
Qy 1501 TGAATAATGTGAGATCTCTGAGCTTATAGGACTGCGCAATCCGAAGACTGACAGAGG 1560
Db 2500 TGAATAATGTGAGATCTCTGAGCTTATAGGACTGCGCAATCCGAAGACTGACAGAGG 2559
Qy 1561 CCCTTTGCTTGGATCTCTTTGAGCCTGTGAGCTGATGATGCTTGGACCGACAGCAAC 1620
Db 2560 CCCTTTGCTTGGATCTCTTTGAGCCTGTGAGCTGATGATGCTTGGACCGACAGCAAC 2619
Qy 1621 TCAAGCAAAATGACAGCCCATGATCTCTGAGATTTAATTTGTTGACCACTATTT 1680
Db 2620 TCAAGCAAAATGACAGCCCATGATCTCTGAGATTTAATTTGTTGACCACTATTT 2679
Qy 1681 ATGACCGCTCTGGACCAAGAGACCAAAATTTGGTCAACGCTCTCTGCGTGGATATGT 1740
Db 2680 ATGACCGCTCTGGACCAAGAGACCAAAATTTGGTCAACGCTCTCTGCGTGGATATGT 2739
Qy 1741 GTCTGAATCTGCTGCTGATGTTTATGATAGGAGCAAGACAGGAGGATCGCTGCTCTGT 1800
Db 2740 GTCTGAATCTGCTGCTGATGTTTATGATAGGAGCAAGACAGGAGGATCGCTGCTCTGT 2799
Qy 1801 CTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAAGTACAGATACC 1860
Db 2800 CTTTAAAACTGGCATCATTTCCCTGTGTAAAGCACATTTGGAAAGACAAAGTACAGATACC 2859
Qy 1861 TTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACCGAGCGAGCTGGGCTCTCTTC 1920
Db 2860 TTTTCAAGCAAGTGGCAAGTTTCAACAGATTTTGTGACCGAGCGAGCTGGGCTCTCTTC 2919
Qy 1921 TGCATGATTTCTATCAAAATCCAAAGACAGTTGGGTGAAGTTGCATCTCTTTGGGGGACAGTA 1980
Db 2920 TGCATGATTTCTATCAAAATCCAAAGACAGTTGGGTGAAGTTGCATCTCTTTGGGGGACAGTA 2979
Qy 1981 ACATTGACCCAAAGTGTCCGGA 2001
Db 2980 ACATTGACCCAAAGTGTCCGGA 3000

RESULT 2

US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1

GENERAL INFORMATION:

APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
TITLE OF INVENTION: THEREOF

FILE REFERENCE: DE1142

CURRENT APPLICATION NUMBER: US/09/845,416

CURRENT FILING DATE: 2001-04-30

PRIOR APPLICATION NUMBER: 60/200,777

PRIOR FILING DATE: 2000-04-28

NUMBER OF SEQ ID NOS: 36

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 32

LENGTH: 4414

TYPE: DNA

ORGANISM: Homo sapiens

US-09-845-416-32

Query Match

99.5%; Score 1990; DB 13; Length 4414;

Best Local Similarity 100.0%; Pred. No. 0;				Matches 2001; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			
QY	1	GGCAGTTTCATTGAGAGAGTGAAGTAACTGACCGTTATCAACAGAGCTTTAGAGAA	60				
DB	1757	GGCAGTTTCATTGAGAGAGTGAAGTAACTGACCGTTATCAACAGAGCTTTAGAGAA	1816				
QY	61	GTATTATCGTGGCTTTCTTTCTGCTGAGAGACATTTGCAAGCACAAGGAGAGATTTCTTAAT	120				
DB	1817	GTATTATCGTGGCTTTCTTTCTGCTGAGAGACATTTGCAAGCACAAGGAGAGATTTCTTAAT	1876				
QY	121	GATGTGGAAGTGGTGAAGAGACAGTTTCATCTCATGAGGGGTACATGATGGATTGACA	180				
DB	1877	GATGTGGAAGTGGTGAAGAGACAGTTTCATCTCATGAGGGGTACATGATGGATTGACA	1936				
QY	181	GCCCATCAGGCGCGGTGGTAAATTTCTCAATTTGGGAAGTGAAGTGGGAACAGGA	240				
DB	1937	GCCCATCAGGCGCGGTGGTAAATTTCTCAATTTGGGAAGTGAAGTGGGAACAGGA	1996				
QY	241	AAATTATCAGGAAGTGAAGAACTGAAGTACAAGAGAGATGAATCTCTAAATTTCAAGA	300				
DB	1997	AAATTATCAGGAAGTGAAGAACTGAAGTACAAGAGAGATGAATCTCTCTAAATTTCAAGA	2056				
QY	301	TGGGAATGCCCTCAGGTTAGTACATGGAAGAAACAAGCAATTTACATAGATTTTAATG	360				
DB	2057	TGGGAATGCCCTCAGGTTAGTACATGGAAGAAACAAGCAATTTACATAGATTTTAATG	2116				
QY	361	GATCTCCAGAAATC-GAAACTGAAAGAGTGTGAATGACTGGCTTAACAAAAACAGAGAAAGA	419				
DB	2117	GATCTCCAGAAATCAGAAATCAGAAAGTGTGAATGACTGGCTTAACAAAAACAGAGAAAGA	2176				
QY	420	ACAAGGAATGGAGAGAGAGCCCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA	479				
DB	2177	ACAAGGAATGGAGAGAGAGCCCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA	2236				
QY	480	CAACAACTAAGTGTCTTCAAGAGAGTCTAGAAACAAGCAAGTCAAGGTCATTTCTCTC	539				
DB	2237	CAACAACTAAGTGTCTTCAAGAGAGTCTAGAAACAAGTCAAGGTCATTTCTCTC	2296				
QY	540	ACTCACATGGTGGTGGTATGTAATCTAGTGGAGATCAGCAACTGTCTGTCTTGGAA	599				
DB	2297	ACTCACATGGTGGTGGTATGTAATCTAGTGGAGATCAGCAACTGTCTGTCTTGGAA	2356				
QY	600	GAAACAATTAGTATGGAGATCGATGGCAACATCTGTAGATGACAGAGACCGC	659				
DB	2357	GAAACAATTAGTATGGAGATCGATGGCAACATCTGTAGATGACAGAGACCGC	2416				
QY	660	TGGGTTCTTTTACAAGACATCCTTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTT	719				
DB	2417	TGGGTTCTTTTACAAGACATCCTTTCTCAAAATGGCAACGTCTTACTGAAGAACAGTGCCTT	2476				
QY	720	TTTAGTGCATGGCTTTCAGAAAAAGAGATGACGTGAACAGATTTACACAACTGGCTTT	779				
DB	2477	TTTAGTGCATGGCTTTCAGAAAAAGAGATGACGTGAACAGATTTACACAACTGGCTTT	2536				
QY	780	AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAATGGCCGTTTAAAGCGGATCTA	839				
DB	2537	AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAATGGCCGTTTAAAGCGGATCTA	2596				
QY	840	GAAAGAAAAAGCAATCCATGGGCAAACTGATTCATCTCAAAACAGATCTTTTCAACA	899				
DB	2597	GAAAGAAAAAGCAATCCATGGGCAAACTGATTCATCTCAAAACAGATCTTTTCAACA	2656				
QY	900	CTGAGATATAGTACGTGACCCAGAGACGAGCATGGCTGGGATACCTTTGCCCGGTGT	959				
DB	2657	CTGAGATATAGTACGTGACCCAGAGACGAGCATGGCTGGGATACCTTTGCCCGGTGT	2716				
QY	960	TGGGTAATTTTAGTCCAAAACTTTGAAAGAGTACAGCACAGACCTTTGAAAGCTCCAG	1019				
DB	2717	TGGGTAATTTTAGTCCAAAACTTTGAAAGAGTACAGCACAGACCTTTGAAAGCTCCAG	2776				
QY	1020	GAACTTCAAGAGCCACCGATGAGCTGAGCTCAAGCTGCGCCCAAGCTGAGGTGATCAAG	1079				

DB	2777	GAACTTCAAGAGCCACCGATGAGCTGAGCTCAAGCTGCGCCAAAGCTGAGGTGATCAAG	2836
QY	1080	GGATCTCTGACAGCCCGTGGCGGATCTCTCTCAATTTGACTCTCTCCAAAGATCACTCCAGAAA	1139
DB	2837	GGATCTCTGACAGCCCGTGGCGGATCTCTCTCAATTTGACTCTCTCCAAAGATCACTCCAGAAA	2896
QY	1140	GTCAAGGCACCTTCGAGAGAGAAATTCGGCTCTGAAAGAGAAAGTGAAGCAAGTCAATGAC	1199
DB	2897	GTCAAGGCACCTTCGAGAGAGAAATTCGGCTCTGAAAGAGAAAGTGAAGCAAGTCAATGAC	2956
QY	1200	CTTGTCTGCGCAGCTTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAAAGCTCAGCACTCTG	1259
DB	2957	CTTGTCTGCGCAGCTTTACCACTTTTGGGCAATTCAGCTCTCACCGTATAAAGCTCAGCACTCTG	3016
QY	1260	GAAAGACCTGAAACACAGATGGAAGCTTTCTGAGAGTGGCGTTCGAGGACCGAGTCAAGGAG	1319
DB	3017	GAAAGACCTGAAACACAGATGGAAGCTTTCTGAGAGTGGCGTTCGAGGACCGAGTCAAGGAG	3076
QY	1320	CTGCATGAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTCTTTTCCACGCTCTGTC	1379
DB	3077	CTGCATGAGCCACAGGAGCTTTGGTCCAGCATCTCAGCACTCTTTTCCACGCTCTGTC	3136
QY	1380	CAGGCTCTCTGAGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACCAG	1439
DB	3137	CAGGCTCTCTGAGAGAGAGCCATCTCGCCAAACAAAGTGCCTTACTATATCAACACCAG	3196
QY	1440	ACTCAAAACAACTTGTGGGACCATCTCCAAAATGACAGAGCTCTACCACTCTTTAGCTGAC	1499
DB	3197	ACTCAAAACAACTTGTGGGACCATCTCCAAAATGACAGAGCTCTACCACTCTTTAGCTGAC	3256
QY	1500	CTGAATATGTCAGATCTCTCAGCTTATAGGACTGCGCACTCCGAAGACTGCAGAGAG	1559
DB	3257	CTGAATATGTCAGATCTCTCAGCTTATAGGACTGCGCACTCCGAAGACTGCAGAGAG	3316
QY	1560	GCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGTGATGCTGTGATGCTGTGATGCTGTGATG	1619
DB	3317	GCCCTTTGCTTGGATCTCTTGGAGCTGTGAGCTGTGATGCTGTGATGCTGTGATGCTGTGATG	3376
QY	1620	CTCAAGCAAAATGACAGCCCATGATATCTGCAAGATTTAATTTGTTGACCACTATT	1679
DB	3377	CTCAAGCAAAATGACAGCCCATGATATCTGCAAGATTTAATTTGTTGACCACTATT	3436
QY	1680	TATGACCGCTGAGAGAGAGACACAACTTTGCTCAACGCTCTCTCTGCTGGATATG	1739
DB	3437	TATGACCGCTGAGAGAGAGACACAACTTTGCTCAACGCTCTCTCTGCTGGATATG	3496
QY	1740	TGTCTGAATCTGCTCTGATGTTTATGATACGGAACGAAAGGAGGATCCGCTGCTG	1799
DB	3497	TGTCTGAATCTGCTCTGATGTTTATGATACGGAACGAAAGGAGGATCCGCTGCTG	3556
QY	1800	TCCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGCAAGTACAGATAC	1859
DB	3557	TCCTTTTAAACCTGGCATCATTTCCCTGTGTAAAGCACATTTTGGAAAGCAAGTACAGATAC	3616
QY	1860	CTTTTCAAGAGTGGAGTTCACAGATTTTGTGACGAGGAGGCTGGGCTCTCTT	1919
DB	3617	CTTTTCAAGAGTGGAGTTCACAGATTTTGTGACGAGGAGGCTGGGCTCTCTT	3676
QY	1920	CTGCATGATCTTATCCAAATTTCCAGAGAGTGGTGAAGTTCATCTTTGGGGGAGT	1979
DB	3677	CTGCATGATCTTATCCAAATTTCCAGAGAGTGGTGAAGTTCATCTTTGGGGGAGT	3736
QY	1980	AAACATTTAGCCCAAGTGTCCGGA 2001	
DB	3737	AAACATTTAGCCCAAGTGTCCGGA 3758	

RESULT 3
US-10-149-736-39
; Sequence 39, Application US/10149736
; Publication No. US2003021632A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.

APPLICANT: Harper, Scott Q.
TITLE OF INVENTION: Mini-Dystrophin Nucleic Acids and Peptide Sequences
FILE REFERENCE: UM-06968
CURRENT APPLICATION NUMBER: US/10/149,736
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: PCT/US01/31126
PRIOR FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,848
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn version 3.1
SEQ ID NO 39
LENGTH: 5417
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Synthetic
US-10-149-736-39

Query Match 91.0%; Score 1820.8; DB 13; Length 5417;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 2; Indels 157; Gaps 2;

Qy 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGAA 60
Db 1199 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTTATCAAAACAGCTTTAGAAAGAA 1258

Qy 61 GTATTATTCGTGGCTCTCTTCTGCTGAGGACACATTCGAAGCAAGAGAGATTTCTTAAT 120
Db 1259 GTATTATTCGTGGCTCTCTTCTGCTGAGGACACATTCGAAGCAAGAGAGATTTCTTAAT 1318

Qy 121 GATGTGAAGTGTGGAAGACCAAGTTCATCTATCATGAGGGGTACATGATGGATTTGACA 180
Db 1319 GATGTGAAGTGTGGAAGACCAAGTTCATCTATCATGAGGGGTACATGATGGATTTGACA 1378

Qy 181 GCCCATCAGGCCCGGTGGTATATTTCTACATTTGGAGTAAAGCTGATGGAAACAGGA 240
Db 1379 GCCCATCAGGCCCGGTGGTATATTTCTACATTTGGAGTAAAGCTGATGGAAACAGGA 1438

Qy 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGA 300
Db 1439 AAATTATCAGAAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTCAGA 1498

Qy 301 TGGGAATGCCCTCAGGTAGCTAGCATGGAACAAAGCAATTTACATAGAGTTTAAATG 360
Db 1499 TGGGAATGCCCTCAGGTAGCTAGCATGGAACAAAGCAATTTACATAGAGTTTAAATG 1558

Qy 361 GATCTCCAGATC-GAAACTGAAAGAGTTGAATGACTGGCTAAACAAAGACAGAAAGAA 419
Db 1559 GATCTCCAGATCAGAACTGAAAGAGTTGAATGACTGGCTAAACAAAGACAGAAAGAA 1618

Qy 420 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGCCTAAACCCCAAGTA 479
Db 1619 ACAAGGAAATGGAGGAGAGCCCTCTTGGACCTGATCTTGAAGCCTAAACCCCAAGTA 1678

Qy 480 CAACAACATAGGTGCTCTCAGAGAGATCTAGAACAGAAAGCAAGTCAAGGGTCAATTTCTC 539
Db 1679 CAACAACATAGGTGCTCTCAGAGAGATCTAGAACAGAAAGCAAGTCAAGGGTCAATTTCTC 1738

Qy 540 ACTCACATGGTGGTGGTATGATGAATCTAGTGGAGATCACGCACTGTGCTTTGGAA 599
Db 1739 ACTCACATGGTGGTGGTATGATGAATCTAGTGGAGATCACGCACTGTGCTTTGGAA 1798

Qy 600 GAACNACTTAGGTATTTGGGAGATCGATGGCAACATCTGTAGATGACAGAGAGCCGC 659
Db 1799 GAACNACTTAGGTATTTGGGAGATCGATGGCAACATCTGTAGATGACAGAGAGCCGC 1858

Qy 660 TGGGTCTCTTTTACAGACATCCTCTCAATGGAAGCAAGCTTTTACTGAAGACAGTGCCTT 719
Db 1859 TGGGTCTCTTTTACAGACATCCTCTCAATGGAAGCAAGCTTTTACTGAAGACAGTGCCTT 1918

Qy 720 TTTAGTGCATGGCTTTCAAGAAAGAGAGTGCAGTGAACAGATTTACACACTGGCTTT 779

Db 1919 TTTAGTGCATGGCTTTCAAGAAAAGAGATGCAGTGAACAAAGATTTACACAACTGGCTTT 1978

Qy 780 AAAGATCAAAATGAATGAAATGTTATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTA 839

Db 1979 AAAGATCAAAATGAATGAAATGTTATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTA 2038

Qy 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCACTCAAAAGATCTTCTTTCAACA 899

Db 2039 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCACTCAAAAGATCTTCTTTCAACA 2098

Qy 900 CTGAAGAAATAGTCACTGAGCCGAGAACGAGCATGGTGGATACACTTTGCCCCGCTGT 959

Db 2099 CTGAAGAAATAGTCACTGAGCCGAGAACGAGCATGGTGGATACACTTTGCCCCGCTGT 2158

Qy 960 TGGGATTAATTTAGTCCAAAACCTTTGAAAAGAGTACAGACAGA----- 1002

Db 2159 TGGGATTAATTTAGTCCAAAACCTTTGAAAAGAGTACAGACAGAATTTTCACAGGCTGTCA 2218

Qy 1003 ----- 1002

Db 2219 ACCACTCAGCCATCACTAAACACAGACAACCTGTATATGGAACAGTAACGTGACCA 2278

Qy 1003 ----- 1002

Db 2279 AGGNAACAGATCTCTGTTAAAGCATGCTCAAGAGGAACTTCCACCACCCTCCCCAAAG 2338

Qy 1003 -----CCCTTGAAGACTCCAGAACTTCAAGAGGCCACGGATGAG 1043

Db 2339 AAGAGGCAAGATTACTGTGGATCTTTGAAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAG 2398

Qy 1044 CTGGACCTCAAGCTGGCCAGCTGAGTGTATCAAGGGATCTCTGGCAGCCGCTGGGCGAT 1103

Db 2399 CTGGACCTCAAGCTGGCCAGCTGAGTGTATCAAGGGATCTCTGGCAGCCGCTGGGCGAT 2458

Qy 1104 CTCCTCATTTGACTCTCTCCAGATCACTCGAGAAAGTCAAGGCACCTCGAGGAGAAATT 1163

Db 2459 CTCCTCATTTGACTCTCTCCAGATCACTCGAGAAAGTCAAGGCACCTCGAGGAGAAATT 2518

Qy 1164 GCGCTCTGAAAAGAGAGTGGAGCCAGTCAATGACCTTCTGCCAGCTTACCACCTTTG 1223

Db 2519 GCGCTCTGAAAAGAGAGTGGAGCCAGTCAATGACCTTCTGCCAGCTTACCACCTTTG 2578

Qy 1224 GGCATTCAAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAG 1293

Db 2579 GGCATTCAAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAAACACAGATGGAAG 2638

Qy 1284 CTTCTGCAAGTGGCCGTGAGGACCGAGTCAAGGAGTGCATGAGGCCACAGGAGCTTT 1343

Db 2639 CTTCTGCAAGTGGCCGTGAGGACCGAGTCAAGGAGTGCATGAGGCCACAGGAGCTTT 2698

Qy 1344 GGTCCAGCATCTCAGCACTTTCTTTTCCACGCTCTGTCAGGGTCCCTGGAGAGAGCCATC 1403

Db 2699 GGTCCAGCATCTCAGCACTTTCTTTTCCACGCTCTGTCAGGGTCCCTGGAGAGAGCCATC 2758

Qy 1404 TGGCCAAAACAAAGTGCCTACTATATCAACCAAGAGTCAAACTGCTGGGAGCCAT 1463

Db 2759 TGGCCAAAACAAAGTGCCTACTATATCAACCAAGAGTCAAACTGCTGGGAGCCAT 2818

Qy 1464 CCCAAAATGACAGAGCTTACAGCTTTTAGCTGACCTGATATATGTCAGATTTCTCAGCT 1523

Db 2819 CCCAAAATGACAGAGCTTACAGCTTTTAGCTGACCTGATATATGTCAGATTTCTCAGCT 2878

Qy 1524 TATAGACTGTCATGAAACTCCGAGAGTGCAGAAAGCCCTTTTCTTGGATCTTTGAGC 1583

Db 2879 TATAGACTGTCATGAAACTCCGAGAGTGCAGAAAGCCCTTTTCTTGGATCTTTGAGC 2938

Qy 1584 CTGTCACTGATGTGATGCTTGGACCCAGCAACCTCAGCAAAATGACACCCCATG 1643

Db 2939 CTGTCACTGATGTGATGCTTGGACCCAGCAACCTCAGCAAAATGACACCCCATG 2998

Qy 1644 GATATCTTCAGATTAATTAATTTGTTGACCACTATTTATGACCCGCTGGAGCAAGAGCAC 1703

Db 2999 GATATCTTCAGATTAATTAATTTGTTGACCACTATTTATGACCCGCTGGAGCAAGAGCAC 3058

QY 1704 AACAAATTTGGTCAACGCTCCCTCTCTGGTGGATATGTCTGAACCTGGCTGCTGAATGTT 1763
Db 3059 AACAAATTTGGTCAACGCTCCCTCTCTGGTGGATATGTCTGAACCTGGCTGCTGAATGTT 3118
QY 1764 TATGATAGGGAGCAACAGGAGAGATCCGTGCTCTGCTCTTTTAAACTGGCATCATTTCC 1823
Db 3119 TATGATAGGGAGCAACAGGAGAGATCCGTGCTCTGCTCTTTTAAACTGGCATCATTTCC 3178
QY 1824 CTGTGTAAGCACATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA 1883
Db 3179 CTGTGTAAGCACATTTTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCA 3238
QY 1884 ACAGATTTTGTGACCGGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCA 1943
Db 3239 ACAGATTTTGTGACCGGAGGCTGGGCTCTCTCTGATGATTTCTATCCAAATTTCCA 3298
QY 1944 AGACAGTTGGGTGAAGTTGCATCTTTGGGGGAGTAACTAGGCAAGTGTCCCGA 2001
Db 3299 AGACAGTTGGGTGAAGTTGCATCTTTGGGGGAGTAACTAGGCAAGTGTCCCGA 3356

RESULT 4
US-09-845-416-12
; Sequence 12, Application US/09845416
; Publication No. US2003017312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-12

Query Match 79.6%; Score 1593; DB 13; Length 3510;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1811; Conservative 0; Mismatches 190; Indels 64; Gaps 4;
QY 1 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGAGCTTTAGAAGAA 60
Db 1000 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGGACCGTTATCAACAGAGCTTTAGAAGAA 1059
QY 61 GTATTATCGTGGCTCTTCTCTGCTGAGACACATTGCAAGCACAGAGAGATTTCTAAT 120
Db 1060 GTATTATCGTGGCTCTTCTCTGCTGAGACACATTGCAAGCACAGAGAGATTTCTAAT 1119
QY 121 GATGTGGAAGTGGTGAAGAACAGTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 180
Db 1120 GATGTGGAAGTGGTGAAGAACAGTTCATCTACTCATGAGGGGTACATGATGGATTTGACA 1179
QY 181 GCCCATCAGGGCCGGTGGTAAATTTCTACAAATGGGAAGTAACTGATGGACAGGA 240
Db 1180 GCCCATCAGGGCCGGTGGTAAATTTCTCAATTTGGGAAGTAACTGATGGACAGGA 1239
QY 241 AAATTTATCAGAAGTGAAGAACTGAAGTACAGAGCAGATGAATCTCTCTAAATTTCAAGA 300
Db 1240 AAATTTATCAGAAGTGAAGAACTGAAGTACAGAGCAGATGAATCTCTCTAAATTTCAAGA 1299
QY 301 TGGGAATGCTCAGGGTATGATGATGGAATAAACAAGCAATTTACATAGATTTTAATG 360
Db 1300 TGGGAATGCTCAGGGTATGATGATGGAATAAACAAGCAATTTACATAGATTTTAATG 1359
QY 361 GATCTCCAGATC-GAAACTGAAGAGTTGAATGATCTGCTTAAACAAAAACAGAGAAAGA 419

Db 1360 GATCTCCAGATCAGAAACTGAAGAGTTGAATGATCTGCTGTAAACAAACAGAGAAAGA 1419
QY 420 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 479
Db 1420 ACAAGGAAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTTGAAGACCTTAAACGCCAAGTA 1479
QY 480 CAAACAACATAGGTGCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGTCAATTTCTCTC 539
Db 1480 CAAACAACATAGGTGCTTCAAGAGATCTAGAACAAAGAACAAAGTCAGGTCAATTTCTCTC 1539
QY 540 ACTCATCATGCTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db 1540 ACTCATCATGCTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 1599
QY 600 GAAACAATTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGACAGAAACCGC 659
Db 1600 GAAACAATTAAGTATTTGGAGATCGATGGGCAACATCTGTAGATGACAGAAACCGC 1659
QY 660 TGGGTTCTTTTACAGACATCCCTCTCAAAATGGCAACGCTTCTACTGAAGAACAGTGCCTT 719
Db 1660 TGGGTTCTTTTACAGACATGTTCTGACCAAGTGGAGAGCTCTGCACCTTCTCTGAGGNA 1719
QY 720 TTTAGTGCATGGCTTTTCAAAAAAGAGATGAGTGAACAAAGATTCACAACTGGCTTTT 779
Db 1720 CTTCTGGTGTGGCTACAGCTGAAGATGATGAATTAAGCGCGCAGGCACCTATTGGAGGC 1779
QY 780 AAGATCAAAATGAATGTTATCAAGTCTTCAAACACTGGCGGTTTTTAAAGCGGATCTA 839
Db 1780 GACTTTCCACAGAGTTCAGAAAGCAGAACGATGTACATAGGGGCTTCAAGAGGGAATTGAAA 1839
QY 840 GAAAAAGAAAAAGCAATCCATGGGCAAACTGTA-----TTCACTCAAAACAG 885
Db 1840 ACTAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAATATTCTGACAGAGCAG 1899
QY 886 ATCTCTTTTCAACACTGAAGATAAGTCAAGTACCCAGCAAGACGGAAGATGGCTGGA--- 943
Db 1900 CCTTTGGAGGACTAGAGAACTCTACCAGGAGCCACAGAGAGCTGCTCTGAGGAGAGA 1959
QY 944 -----TAACTTTGCCCGGTGTTGGGATAAATTAGTCCAAA 978
Db 1960 GCCCAGATGTCACTCGGCTTCTACGAAGACAGGCTGAGGAGGTCAATCTAGTGGGAA 2019
QY 979 AACTTGAA-----AAGATACAGCACAGACCTTTGAAAGACTC 1016
Db 2020 AAATTTGAACCTGCACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAAGACTC 2079
QY 1017 CAGGAACTTCAAGAGCCACGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATC 1076
Db 2080 CAGGAACTTCAAGAGCCACGATGAGTGGACCTCAAGCTGCGCCAAAGCTGAGGTGATC 2139
QY 1077 AAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAG 1136
Db 2140 AAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTTGACTCTCTCCAAGATCACTCGAG 2199
QY 1137 AAGTCAAGGCACTTCGAGGAGAAATTTGGGCTCTGAAAGAGAACGTGAGCCACGTCAAT 1196
Db 2200 AAGTCAAGGCACTTCGAGGAGAAATTTGGGCTCTGAAAGAGAACGTGAGCCACGTCAAT 2259
QY 1197 GACCTTGTCTGCGCAGCTTACCACTTTGGGCATTTCAGCTCTCACCGTATAAACCCTCAGACT 1256
Db 2260 GACCTTGTCTGCGCAGCTTACCACTTTGGGCATTTCAGCTCTCACCGTATAAACCCTCAGACT 2319
QY 1257 CTGGAAGACCTGAAACACAGATGGAAGCTTTGAGGTGGCGGTGAGGACCGAGTCAAG 1316
Db 2320 CTGGAAGACCTGAAACACAGATGGAAGCTTTGAGGTGGCGGTGAGGACCGAGTCAAG 2379
QY 1317 CAGCTGCATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTCT 1376
Db 2380 CAGCTGCATGAAGCCACAGGCACTTTGGTCCAGCATCTCAGCACTTTCTTCCACGCTCT 2439
QY 1377 GTCCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCCTACTATATCAACAC 1436
Db 2440 GTCCAGGGTCCCTGGGAGAGACCATCTCGCCAAACAAAGTGCCCTACTATATCAACAC 2499

```
QY 1437 GAGACTCAACAACTTGGTGGGACCATCCAAAATGACAGAGCTCTACAGCTTTTGTCT 1496
Db 2500 GAGACTCAACAACTTGGTGGGACCATCCAAAATGACAGAGCTCTACAGCTTTTGTCT 2559
QY 1497 GACCTGAATAATCTCAGATCTCAGCTTATAGAGTCCCATGAAATCCGAAAGATCGAG 1556
Db 2560 GACCTGAATAATCTCAGATCTCAGCTTATAGAGTCCCATGAAATCCGAAAGATCGAG 2619
QY 1557 ARGCCCTTTGCTTGGATCTTGGACCTGTGAGCTGTGATGATGATGATGATGATGAT 1616
Db 2620 ARGCCCTTTGCTTGGATCTTGGACCTGTGAGCTGTGATGATGATGATGATGATGAT 2679
QY 1617 AACCTCAAGCAAAATGACCAAGCCATGGATATCTCGAGATTTAATTTGTTGACCACT 1676
Db 2680 AACCTCAAGCAAAATGACCAAGCCATGGATATCTCGAGATTTAATTTGTTGACCACT 2739
QY 1677 ATTATGACCGCTGGAGCAAGAGACACAATTTGGTCAACGCTCCTCTCGCGTGTAT 1736
Db 2740 ATTATGACCGCTGGAGCAAGAGACACAATTTGGTCAACGCTCCTCTCGCGTGTAT 2799
QY 1737 ATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 1796
Db 2800 ATGTGCTGAACCTGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTC 2859
QY 1797 CTGTCTTTTAAATCTGCATCATTTCCCTGTGTAAAGACATTTTGAAGACAGATACAGA 1856
Db 2860 CTGTCTTTTAAATCTGCATCATTTCCCTGTGTAAAGACATTTTGAAGACAGATACAGA 2919
QY 1857 TACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCGAGCGAGGCTGGGCCCTC 1916
Db 2920 TACCTTTTCAAGCAAGTGGCAAGTTTCAACAGGATTTTGTACCGAGCGAGGCTGGGCCCTC 2979
QY 1917 CTTCTGCATGATTTATCCAAATTTCCAAGACAGTTGGTGAAGTTGCATCTCTTTGGGGGC 1976
Db 2980 CTTCTGCATGATTTATCCAAATTTCCAAGACAGTTGGTGAAGTTGCATCTCTTTGGGGGC 3039
QY 1977 AGTAACATTGAGCAAGTGTCCGGA 2001
Db 3040 AGTAACATTGAGCAAGTGTCCGGA 3064

RESULT 5
US-09-845-416-31
; Sequence 31, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 31
; LENGTH: 4476
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-31

Query Match 79.6%; Score 1593; DB 13; Length 4476;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1811; Conservative 0; Mismatches 190; Indels 64; Gaps 4;

QY 1 GGCAGTTCATTGATGAGAGTGAAGTAACTCGACCGTTATCAAAACAGCTTTAGAGAA 60
Db 1756 GGCAGTTCATTGATGAGAGTGAAGTAACTCGACCGTTATCAAAACAGCTTTAGAGAA 1815
QY 61 GTATTATCGTGGCTCTTCTGCTGAGACACATTTGCAAGCACAGGAGAGATTTCTTAAT 120
```

```
Db 1816 GTATTATCGTGGCTCTTCTGCTGAGGACACATTTGCAAGCACAGGAGAGATTTCTTAAT 1875
QY 121 GATGTGGAAGTGGTGAAGACCACTTTTCTACTCATGAGCGGTACATGATGGATTTTGACA 180
Db 1876 GATGTGGAAGTGGTGAAGACCACTTTTCTACTCATGAGCGGTACATGATGGATTTTGACA 1935
QY 181 GCCCATCAGGCGCGGGTGTGTAATATTTCTACAATTTGGGAAAGTAAAGCTGATTTGNAACAGGA 240
Db 1936 GCCCATCAGGCGCGGGTGTGTAATATTTCTACAATTTGGGAAAGTAAAGCTGATTTGNAACAGGA 1995
QY 241 AAATTTATCAGAAGATGAAGAAACTGGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 300
Db 1996 AAATTTATCAGAAGATGAAGAAACTGGAAGTACAAGAGCAGATGAATCTCTTAATTTCAAGA 2055
QY 301 TGGGAATGCTCTCAGGCTAGCTAGCATGGAAAAAACAAGCAATTTTACATAGATTTTAATG 360
Db 2056 TGGGAATGCTCTCAGGCTAGCTAGCATGGAAAAAACAAGCAATTTTACATAGATTTTAATG 2115
QY 361 GATCTCCAGATC-GAAACTTGAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGA 419
Db 2116 GATCTCCAGATCAGAAACTGAAAGAGTTGAATGACTGGCTTAACAAAAACAGAGAAAGA 2175
QY 420 ACAAGGAAAAATGAGGAGAGAGCCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTA 479
Db 2176 ACAAGGAAAAATGAGGAGAGAGCCCTCTTTGGACCTGATCTTGAAGACCTTAAACCGCAAGTA 2235
QY 480 CAACAAATTAAGTGTCTTCAAGAAAGTCTAGAACAGAAACAAGTCAGGTCATTTCTCTC 539
Db 2236 CAACAAATTAAGTGTCTTCAAGAAAGTCTAGAACAGAAACAAGTCAGGTCATTTCTCTC 2295
QY 540 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 599
Db 2296 ACTCACATGGTGGTGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAA 2355
QY 600 GAAACAATTAAGTGTGGAGATCGATGGGCAAAACATCTGTAGATGACAGAGACCGC 659
Db 2356 GAAACAATTAAGTGTGGAGATCGATGGGCAAAACATCTGTAGATGACAGAGACCGC 2415
QY 660 TGGGTTCTTTTACAGACATCCTTCTCAAAATGGCAAGTCTTACTGAAGAACAGTGCCTT 719
Db 2416 TGGGTTCTTTTACAGACATCTTCAACAGACATCTTCAACAGTGGAGGCTGTGACCTTCTCTCAGGAA 2475
QY 720 TTTAGTGCATGCTTTTCAAAAAAGAGATGCAAGTGAACAGATTCACAACTGGCTTT 779
Db 2476 CTTCTGGTGTGGCTACAGCTGAAAGATGATGAATTAAGCCGCGCAGGACCCTATTGGAGGC 2535
QY 780 AAAGATCAAAATGAATGTTATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTA 839
Db 2536 GACTTTCCAGCAGTTTCAAGACAGAACGATGTATAGGGCTTCAAGAGGGGAATTTGAA 2595
QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTGA-----TTCACTCAAAACAG 885
Db 2596 ACTAAGAACCTGTAAATCATGAGTACTCTTGAAGACTGTACGAATATTTCTGACAGAGCAG 2655
QY 886 ATCTTTCTTTCAACACTGAAGATTAAGTCAAGTCAAGCCGAGAGACGGAAGCATGGCTGA-- 943
Db 2656 CCTTTGGAAGGACTAGAGAAACTCTACAGAGAGCCAGAGAGCTGCCTCTTGAAGAGAGA 2715
QY 944 -----TAACTTTGCCCGTGTGGGTAATAATTTAGTCCAAA 978
Db 2716 GCCAGAAATGTCACCTCGGCTTCTACGAAGACAGGCTGAGGAGTCAATCTAGTGGGAA 2775
QY 979 AACTTTGAA-----AAGAGTACAGCACAGACCCCTTGAAGACTC 1016
Db 2776 AAATTTGAACCTTGCACCTCGCTGACTGGCAGAGAAAAATAGATGAGACCCCTTGAAGACTC 2835
QY 1017 CAGGAACCTTCAAGAGCCACGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATC 1076
Db 2836 CAGGAACCTTCAAGAGCCACGATGAGCTGAGCTCAAGCTGCGCAAGCTGAGGTGATC 2895
QY 1077 AAGGGATCTGSCAGCCCGTGGGGATCTCTCTAATGATCTCTTCCCAAGATCACTCGAG 1136
Db 2896 AAGGGATCTGSCAGCCCGTGGGGATCTCTCTAATGATCTCTTCCCAAGATCACTCGAG 2955
```

QY	1137	AAAGTCAAGGCACTTCGAGGAGAAAATTGGCCCTCTGAAAGAGAACTGAGCCACGTCAAT	1196
Db	2956	AAAGTCAAGGCACTTCGAGGAGAAATTGGCCCTCTGAAAGAGAACTGAGCCACGTCAAT	3015
QY	1197	GACCTTGCTCGGCAGCTTACCACTTTGGGCATTACGCTCTCACGGTATACCTTCAGCACT	1256
Db	3016	GACCTTGCTCGGCAGCTTACCACTTTGGGCATTACGCTCTCACGGTATACCTTCAGCACT	3075
QY	1257	CTGGAAGACTGGAACACACAGATGAAGCTTCTCGAGTGGCCGTGAGAGCCAGTCAAG	1316
Db	3076	CTGGAAGACTGGAACACACAGATGAAGCTTCTCGAGTGGCCGTGAGAGCCAGTCAAG	3135
QY	1317	CAGCTCATGAAGCCACACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCT	1376
Db	3136	CAGCTCATGAAGCCACACAGGAGCTTTGGTCCAGCATCTCAGCACTTTCTTTCCACGTCT	3195
QY	1377	GTCGAGGTCCCTGGGAGAGAGCCATCTCGCCAAAAPAGTGGCTTACTATATCAACCAC	1436
Db	3196	GTCGAGGTCCCTGGGAGAGAGCCATCTCGCCAAAAPAGTGGCTTACTATATCAACCAC	3255
QY	1437	GAGACTCAAAACAATTTGCTGGGACCACTCCCAAATGACAGAGCTCTACCACTCTTTAGCT	1496
Db	3256	GAGACTCAAAACAATTTGCTGGGACCACTCCCAAATGACAGAGCTCTACCACTCTTTAGCT	3315
QY	1497	GACCTGAATAATGTGAGATTCTCAGCTTATAGACTGGCCATGAAAACTCGAAGACTGCAG	1556
Db	3316	GACCTGAATAATGTGAGATTCTCAGCTTATAGACTGGCCATGAAAACTCGAAGACTGCAG	3375
QY	1557	AAGGCCCTTTGCTTGGATCTCTTGAGCCGTGTCAGCTGCGATGTCATGCTTTGGACCAAGCAC	1616
Db	3376	AAGGCCCTTTGCTTGGATCTCTTGAGCCGTGTCAGCTGCGATGTCATGCTTTGGACCAAGCAC	3435
QY	1617	AACCTCAAGCAAAATGACACGCCCATGATATCCTCGAGATTATTAATTTTGACCACT	1676
Db	3436	AACCTCAAGCAAAATGACACGCCCATGATATCCTCGAGATTATTAATTTTGACCACT	3495
QY	1677	ATTATTAGCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCGCTGGAT	1736
Db	3496	ATTATTAGCCGCTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCGCTGGAT	3555
QY	1737	ATGTGTCGAACGTGGCTGTGAATGTTTATGATACGGGACGAACAGGGAGATCCGTGTC	1796
Db	3556	ATGTGTCGAACGTGGCTGTGAATGTTTATGATACGGGACGAACAGGGAGATCCGTGTC	3615
QY	1797	CTGTCTTTTAAACTGGCATCAATTCCTCTGTAAAGCACATTTTGAAGACAAGTACAGA	1856
Db	3616	CTGTCTTTTAAACTGGCATCAATTCCTCTGTAAAGCACATTTTGAAGACAAGTACAGA	3675
QY	1857	TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCCTC	1916
Db	3676	TACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACGAGCGCAGGCTGGGCCTC	3735
QY	1917	CTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCTCTTTGGGGC	1976
Db	3736	CTTCTGCATGATTCTATCCAAATTCGAAGACAGTTGGGTGAAGTTGCATCTCTTTGGGGC	3795
QY	1977	AGTAACATTGAGCCAAGTGTCCGA	2001
Db	3796	AGTAACATTGAGCCAAGTGTCCGA	3820

RESULT 6

```

RESOLUT 6
US-09-845-416-9
; Sequence 9, Application US/09845416
; Publication NO. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30

```

Query Match	63.1%	Score 1263	DB 13	Length 3858
Best Local Similarity	75.3%	Pred. No. 0		
Matches 1816	Conservative	0	Mismatches 185	Indels 412
Gaps	3			
1	GCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA	60		
1000	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACGCTTTAGAGAA	1059		
61	GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAAT	120		
1060	GTATTATCGTGGCTTTCTTCTGCTGAGGACACATTTGCAAGCACAAGGAGAGATTTCTAAT	1119		
121	GATGTGGAAGTGGTGAAGACACAGTTTCATCTCATGAGGGGTACATGATGATTTGACA	180		
1120	GATGTGGAAGTGGTGAAGACACAGTTTCATCTCATGAGGGGTACATGATGATTTGACA	1179		
181	GCCCATCAGGGCGGGTTCGTTAATATTTCTACAATTTGGGAAGTAAAGCTGATTTGAAACAGGA	240		
1180	GCCCATCAGGGCGGGTTCGTTAATATTTCTACAATTTGGGAAGTAAAGCTGATTTGAAACAGGA	1239		
241	AAATTATCAGAAGATGAAGAACTGAAGTACAAGACGATGATCTCTTAATTTCAAGA	300		
1240	AAATTATCAGAAGATGAAGAACTGAAGTACAAGACGATGATCTCTTAATTTCAAGA	1299		
301	TGGGAATGCTCAGGCTAGCTAGCTAGGAAACCAAGCAATTTACATAGAGTTTAAATG	360		
1300	TGGGAATGCTCAGGCTAGCTAGCTAGGAAACCAAGCAATTTACATAGAGTTTAAATG	1359		
361	GATCTCCAGAATC-GAAACTGAAAGAGTTTGAATGACTGGCTTAACAAAAACAGAGAAAGA	419		
1360	GATCTCCAGAATCAGAAACTGAAAGAGTTTGAATGACTGGCTTAACAAAAACAGAGAAAGA	1419		
420	ACNAGAAAATGGAGGAAGAGCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA	479		
1420	ACAAGAAAATGGAGGAAGAGCTCTTGGACCTGATCTTGAAGACCTTAAACGCCAAGTA	1479		
480	CAACACATAAGTGTCTTCAAGAAGATCTAGAACCAAGAACAAAGTCAAGTCTCTC	539		
1480	CAACACATAAGTGTCTTCAAGAAGATCTAGAACCAAGAACAAAGTCAAGTCTCTC	1539		
540	ACTCATCGTGGTGGTATTGATGATCTAGTGAGATCAAGCAACTGCTGCTTTGGAA	599		
1540	ACTCATCGTGGTGGTATTGATGATCTAGTGAGATCAAGCAACTGCTGCTTTGGAA	1599		
600	GAACACATTTAAGGATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCCG	659		
1600	GAACACATTTAAGGATTGGGAGATCGATGGGCAACATCTGTAGATGGACAGAGACCCG	1659		
660	TGGGTTCTTTTACAAGACATCTCTCAAAATGGCAAGCTCTTACTGAAGAAACAGTGCCTT	719		
1660	TGGGTTCTTTTACAAGACATCTCTCAAAATGGCAAGCTCTTACTGAAGAAACAGTGCCTT	1719		
720	TTTAGTGCATGCTTTTCAGAAAAGNAGATGCAGTGAACAGATTCACAACTGCTTT	779		
1720	TTTCTTGCCTGGCTTACAGAAAGCTGAAAACAACTGCCAATGTCTCTACAGATGCTACCCGT	1779		
780	A-----AAAGTCAAAATGAAATGTTATCAAGTCTTCAAAAACATGGCCGTT	824		
1780	AAGGAAAGGCTCTCAAGACCTCCAGGGAGTAAAGAGCTGATGAAACAAATGGCAAGAC	1839		
825	TTAAAACGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTATTCTACTCAACAA	884		
1840	CTCAAGGTGAAATGAACTCAGACAGATGTTTATCAAACTGTGATGAAACAGCCAA	1899		

QY	885	GATCTTCTTTCAACACTGAAGAAATAAGTCAGTGA	944
Db	1900	AAAACTCTGAGATCCCTGGAAGTCCGATGTCAGTCC	1959
QY	945	AACTTTGCCCGGTGTGGGATAATTTAGTCCAAAAAC	999
Db	1960	AAACATGAACCTCAAGTGGAGTGAACCTTCGAAAAA	2019
QY	1000	-----	999
Db	2020	GAAGCCAGTTCTGACCACTGGAAGCGTCTGCACCT	2079
QY	1000	-----	999
Db	2080	CTACAGCTGAAAGATGATGAATTAACGGCAGGCACC	2139
QY	1000	-----	999
Db	2140	GTTTCAAGCAGAAACGATGTACATAGGGCGTTCAAG	2199
QY	1000	-----	999
Db	2200	GTAATCATGAGTACTCTTGAGACTGTACGAATAT	2259
QY	1000	-----	999
Db	2260	CTAGAGAAACTTACCAGGAGCCACAGAGAGCTGC	2319
QY	1000	-----	999
Db	2320	ACTCGGCTTCTACGAAGCAGGCTGAGGAGGTCAA	2379
QY	1000	-----AGACCCCTTGAAGAGACTCCAGGAAC	1028
Db	2380	CACCTCCGCTGACTGGCAGAGAAAAATAGATGAG	2439
QY	1029	GAGGCCACGGATGAGCTGGAACCTCAAGCTCGGCA	1088
Db	2440	GAGGCCACGGATGAGCTGGACCTCAAGCTCGGCA	2499
QY	1089	CAGCCCGTGGCGATCTCCTCATTTGACTCTCTCA	1148
Db	2500	CAGCCCGTGGCGATCTCCTCATTTGACTCTCTCA	2559
QY	1149	CTTCGAGGAGAAATTTGGCGCTCTGAAGAAGAA	1208
Db	2560	CTTCGAGGAGAAATTTGGCGCTCTGAAGAAGAA	2619
QY	1209	CAGCTTACCACTTTGGCAATTCAGCTCTCAACCG	1268
Db	2620	CAGCTTACCACTTTGGCAATTCAGCTCTCAACCG	2679
QY	1269	AACACACAGATGGAAGCTTTCTGACAGTGGCCGT	1328
Db	2680	AACACACAGATGGAAGCTTTCTGACAGTGGCCGT	2739
QY	1329	GCCACACAGGACATTTGGTCCAGCATCTCAGCACT	1388
Db	2740	GCCACACAGGACATTTGGTCCAGCATCTCAGCACT	2799
QY	1389	TGGGAGAGAGCCATCTTCGCCMAACAAAGTGGCT	1448
Db	2800	TGGGAGAGAGCCATCTTCGCCMAACAAAGTGGCT	2859
QY	1449	ACTTGTCTGGACCATCCCAAAATGACAGAGCTCT	1508
Db	2860	ACTTGTCTGGACCATCCCAAAATGACAGAGCTCT	2919
QY	1509	GTCAGATTTCTCAGCTTTATAGAGCTGCCATGAA	1568
Db	2920	GTCAGATTTCTCAGCTTTATAGAGCTGCCATGAA	2979

RESULT 7

RESOLUTION
US-09-845-416-29

US-09-843-416-23
; Sequence 29, Application US/09845416

; sequence 29, Application US/038
; Publication No. US20030171312A1; PUBLICATION NO: US20
: GENERAL INFORMATION:

; GENERAL INFORMATION:
: APPLICANT: XIAO, XIAO

APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE

;	TITLE OF INVENTION:	DNA SEQU	WHERE SEQU

; TITLE OF INVENTION: THEREOF

FILE REFERENCE: DE1142

; CURRENT APPLICATION NUMBER: US/O

; CURRENT FILING DATE: 2001-04-30

7 CURRENT FILING DATE: 2001-01-20
; PRIOR APPLICATION NUMBER: 60/200,777

;; PRIOR APPLICATION NUMBER:
;; PRIOR FILING DATE: 2000-

; PRIOR FILING DATE: 2000-04-
 ; NUMBER OF SEQ ID NOS: 36

```

; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatattIn Ver 2.1

```

```

; SOFTWARE: Pat
; GEO ID NO 36

```

; SEQ ID NO 29

; LENGTH: 4825

; TYPE: DNA

Query Match 63.1%; Score 1263; DB 13; Length 4825;

Query Match	63.1%	Score	12637
Best Local Similarity	75.3%	Pred.	No. 0:

Best Local Similarity	75.3%;	Pred. No. 0;	Indels	412;	Gaps	3;
Matches 1916;	Conservative	0;	Mismatches	185;		

Qy	1	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA	60
Db	1757	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA	1816
Qy	61	GTATTATCGTGGCTCTTCTTCTGCTGAGGACACATTCGACGACACAAGGAGAGATTTCTTAAT	120
Db	1817	GTATTATCGTGGCTCTTCTTCTGCTGAGGACACATTCGACGACACAAGGAGAGATTTCTTAAT	1876
Qy	121	GATGTGAAGTGTGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA	180
Db	1877	GATGTGAAGTGTGTGAAGACCAAGTTTCATCTCATGAGGGGTACATGATGGATTTGACA	1936
Qy	181	GCCCATCAGGCGCGGGTGGTAAATATTCTTCAATTTGGGAAAGTAAGCTGATTTGAAACAGGA	240
Db	1937	GCCCATCAGGCGCGGGTGGTAAATATTCTTCAATTTGGGAAAGTAAGCTGATTTGAAACAGGA	1996

Db 3520 GCCCAGAGGACTTTGGTCCAGCATCTCAGCACATTTCTTCCACGTCTGTCACGGGTCCT 3579
Qy 1389 TGGGAGAGGACATCTCGCCAAAACAAAGTGCCTACTATATCAACACGAGACTCAACA 1448
Db 3580 TGGGAGAGGACATCTCGCCAAAACAAAGTGCCTACTATATCAACACGAGACTCAACA 3639
Qy 1449 ACTTGTCTGGAGACATCCCAAATGACAGAGCTTACAGACTTTTAGTGTACCTGAATPAAT 1508
Db 3640 ACTTGTCTGGAGACATCCCAAATGACAGAGCTTACAGACTTTTAGTGTACCTGAATPAAT 3699
Qy 1509 GTCAATCTCAGCTTATAGACTGCCATGAATCCCGAAGCTGCAAGAGGCCCTTTGC 1568
Db 3700 GTCAATCTCAGCTTATAGACTGCCATGAATCCCGAAGCTGCAAGAGGCCCTTTGC 3759
Qy 1569 TTGATCTCTTGTAGCTGTACAGTGTGATGCCCTGGACAGCAACACCTCAAGCAA 1628
Db 3760 TTGATCTCTTGTAGCTGTACAGTGTGATGCCCTGGACAGCAACACCTCAAGCAA 3819
Qy 1629 AATGACAGCCCATGGATATCCTGCAGATTAATTAATTTGTTGACCACTATTTATGACCGC 1688
Db 3820 AATGACAGCCCATGGATATCCTGCAGATTAATTAATTTGTTGACCACTATTTATGACCGC 3879
Qy 1689 CTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTCGGTGGATATGTCTCTGAAC 1748
Db 3880 CTGGAGCAAGAGCAACAATTTGGTCAACGTCCCTCTCTCGGTGGATATGTCTCTGAAC 3939
Qy 1749 TGGCTGTGAATTTATGATACGGGAGCAACAGGGAGATCCGTGCTCTGTCTTTAAA 1808
Db 3940 TGGCTGTGAATTTATGATACGGGAGCAACAGGGAGATCCGTGCTCTGTCTTTAAA 3999
Qy 1809 ACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAATTTGGAAGCAATTTGGAAG 1868
Db 4000 ACTGGCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAATTTGGAAGCAATTTGGAAG 4059
Qy 1869 CAATGGCAAGTCAACAGGATTTGTGAACAGCGAGCTGGGCTCTCTCTGCAATGAT 1928
Db 4060 CAATGGCAAGTCAACAGGATTTGTGAACAGCGAGCTGGGCTCTCTCTGCAATGAT 4119
Qy 1929 TCTATCCAAATTCACAGACATTTGGGTGAAGTTCATCTTTGGGCGCAGTAACATGAG 1988
Db 4120 TCTATCCAAATTCACAGACATTTGGGTGAAGTTCATCTTTGGGCGCAGTAACATGAG 4179
Qy 1989 CCAAGTGTCCGGA 2001
Db 4180 CCAAGTGTCCGGA 4192

RESULT 9

US-09-845-416-36
; Sequence 36, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DB1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 36
; LENGTH: 5060
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-36

Query Match 63.1%; Score 1263; DB 13; Length 5060;
Best Local Similarity 75.3%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 185; Indels 412; Gaps 3;

Qy 1 GCAGTTCATTGATGAGAGTGAAGTAAACCTTGGACCGTTATCAAAACAGCTTTTGAAGAA 60
Db 1992 GCAGTTCATTGATGAGAGTGAAGTAAACCTTGGACCGTTATCAAAACAGCTTTTGAAGAA 2051
Qy 61 GTATTATCGTGTCTTCTCTGAGGACACATTTGCAAGCAACAGGAGAGATTTCTAAT 120
Db 2052 GTATTATCGTGTCTTCTCTGAGGACACATTTGCAAGCAACAGGAGAGATTTCTAAT 2111
Qy 121 GATGTGGAAGTGGTGAAGACAGCTTTTCACTATCTATGAGGGGTACATGATGATTTGACA 180
Db 2112 GATGTGGAAGTGGTGAAGACAGCTTTTCACTATCTATGAGGGGTACATGATGATTTGACA 2171
Qy 181 GCCCATCAGGCGCGGTGGTGAATTTCTAATTTGGGAAGTAACTGATTTGGAACAGGA 240
Db 2172 GCCCATCAGGCGCGGTGGTGAATTTCTAATTTGGGAAGTAACTGATTTGGAACAGGA 2231
Qy 241 AAATATCAGAAAGATGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAATTTCAAGA 300
Db 2232 AAATATCAGAAAGTGAAGAACTGAAGTCAAGAGCAGATGAATCTCTTAATTTCAAGA 2291
Qy 301 TGGGAATGCTCAGGGTGTAGCATGGAAGAAACAAAGCAATTTTACATAGAGTTTAAATG 360
Db 2292 TGGGAATGCTCAGGGTGTAGCATGGAAGAAACAAAGCAATTTTACATAGAGTTTAAATG 2351
Qy 361 GATCTCCAGAAATC-CAAACTGGAAGTGAATGATGCTTAACAAACAGGAGAGAAAGA 419
Db 2352 GATCTCCAGAAATCAGAACTGAAGAGTGAATGCTTAACAAACAGGAGAGAAAGA 2411
Qy 420 ACAAGAAATGAGGAAAGAGCTCTTGGACCTGTATCTTTGAAGACCTTAAACGCCAAGTA 479
Db 2412 ACAAGAAATGAGGAAAGAGCTCTTGGACCTGTATCTTTGAAGACCTTAAACGCCAAGTA 2471
Qy 480 CAACACATTAAGTCTTCAAGAGATCTAGAACAGAAACAGTCAAGTCAATTTCTCTC 539
Db 2472 CAACACATTAAGTCTTCAAGAGATCTAGAACAGAAACAGTCAAGTCAATTTCTCTC 2531
Qy 540 ACTCATCTGT 599
Db 2532 ACTCATCTGT 2591
Qy 600 GAAACATTAAGTGTATGAGGAGATGATGGGAAACATCTGTATGATGAGAGAGACCGC 659
Db 2592 GAAACATTAAGTGTATGAGGAGATGATGGGAAACATCTGTATGATGAGAGAGACCGC 2651
Qy 660 TGGGTCTTTTACAGACATCTCTCAATGGCAAGCTTTTACTGGAAGACAGTGCCTT 719
Db 2652 TGGGTCTTTTACAGACATCTCTCAATGGCAAGCTTTTACTGGAAGACAGTGCCTT 2711
Qy 720 TTTAGTGCATGCTTTTACAGAAAGAGATGCAAGTGAACAAAGATTCACAACTGGCTTT 779
Db 2712 TTTAGTGCATGCTTTTACAGAAAGTGAACAAAGTGAACAAAGATTCACAACTGGCTTT 2771
Qy 780 A-----AGATCAAAATGAATGTATCAAGTCTTCAAAACTGGCGTT 824
Db 2772 AAGAAAGCTCTTAGAAGACTCCAAAGGAGTAAAGAGCTGATGAACAAATGCAAGAC 2831
Qy 825 TTAAGGCGGATCTAGAAAGAAAGCAATCCATGGCAAACTGTATTTCACTCAAAACAA 884
Db 2832 CTCCAAAGTGAATTTGAAGCTCACACAGATGTTTATCAACACCTGGATGAAACAGCCAA 2891
Qy 885 GATCTTTTCAACACTGAAGAAATAGTCAAGTACCAGAGAGCGGAGAGATGCTGGAT 944
Db 2892 AAAATCTTGAGATCCCTGGAAAGGTTCCGATGATGACGTCTCTGTTTACAAAGACGTTTGGAT 2951
Qy 945 AACTTTGCCCGGTGTGGGATAATTTAGTCCAAACCTTGAAGAGAGTACAGCAC----- 999
Db 2952 AACATGAACCTCAAGTGGAGTGAACCTTCGGAAGAAAGTCTCTCAACATTTAGTCCCATTTG 3011
Qy 1000 ----- 999
Db 3012 GAAGCCAGTCTTGACCAAGTGAAGCGTCTGCACCTTTCTCTGAGGAACTTCTGCTGGTGG 3071
Qy 1000 ----- 999

Db 3072 CTACAGTGAAGATGATGAATTAAGCGCGCAGGCACCTATTGGAGCGGACTTTCACGCA 3131
QY 1000 ----- 999
Db 3132 GTTCAGAGCAGAACGATGTACATAGGGCCCTTCAGAGGGAAATTGAATAAAGAACTT 3191
QY 1000 ----- 999
Db 3192 GTAAATCATGAGTACTCTTGAGACTGTACGAATATTTCTGACAGAGCAGCGCTTTGGNAGGA 3251
QY 1000 ----- 999
Db 3252 CTAGAGAACTCTACAGAGAGCCAGAGAGCTGCTCTCTGAGGAGAGAGCCAGCAATGTC 3311
QY 1000 ----- 999
Db 3312 ACTCGGCTTCTACGAAGCAGGCTGAGGAGTCAATCTAGTGGGAAAAATTGAACCTG 3371
QY 1000 -----AGACCTTTGAAGACTCCAGGAACCTCAA 1028
Db 3372 CACTCCGCTGACTGGCAGAGAAAAATAGATGAGACCTTTGAAGACTCCAGGAACCTCAA 3431
QY 1029 GAGGCCAGGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGATCTCGG 1088
Db 3432 GAGGCCAGGATGAGTGGACCTCAAGCTGCGCAAGCTGAGGTGATCAAGGATCTCGG 3491
QY 1089 CAGCCCTGGGCGATCTCTCAATGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCA 1148
Db 3492 CAGCCCTGGGCGATCTCTCAATGACTCTCTCAAGATCACCTCGAGAAAGTCAAGGCA 3551
QY 1149 CTTGAGGAGAAATTGGCCCTCTGAAAGAGAAAGTGAGCCAGTCAATGACCTTGTCTGC 1208
Db 3552 CTTGAGGAGAAATTGGCCCTCTGAAAGAGAAAGTGAGCCAGTCAATGACCTTGTCTGC 3611
QY 1209 CAGCTTACCATTGGGCACTTCAAGCTCTCAAGTATTAACCTCAGCACTCTGGAAGACCTG 1268
Db 3612 CAGCTTACCATTGGGCACTTCAAGCTCTCAAGTATTAACCTCAGCACTCTGGAAGACCTG 3671
QY 1269 AACACAGATGAAGCTTCTGAGGTGGCCCTCGAGGACCGAGTCAAGGAGCTGCAATGAA 1328
Db 3672 AACACAGATGAAGCTTCTGAGGTGGCCCTCGAGGACCGAGTCAAGGAGCTGCAATGAA 3731
QY 1329 GCGCAGAGGACTTGGTCCAGCATCTCAGCACTTCTTCCAGCTCTGCCAGGCTCC 1388
Db 3732 GCGCAGAGGACTTGGTCCAGCATCTCAGCACTTCTTCCAGCTCTGCCAGGCTCC 3791
QY 1389 TGGGAGAGGCACTCTCGCCAAAACAAAGTGGCCCTACTATATCAACACAGAGACTCAAACA 1448
Db 3792 TGGGAGAGGCACTCTGCGCAACAAAGTGGCCCTACTATATCAACACAGAGACTCAAACA 3851
QY 1449 ACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAT 1508
Db 3852 ACTTGTGGGACCAATCCCAAAATGACAGAGCTCTACAGTCTTTAGCTGACCTGAATAT 3911
QY 1509 GTGAGATTCTCAGCTTTAGAGCTGCCATGAACCTCCGAAGACTGCAAGAGGCCCTTTGC 1568
Db 3912 GTGAGATTCTCAGCTTTAGAGCTGCCATGAACCTCCGAAGACTGCAAGAGGCCCTTTGC 3971
QY 1569 TTGATCTCTTGGCTGTGAGCTGCAATGATGCCCTTGGACAGCAACACCTCAAGCAA 1628
Db 3972 TTGATCTCTTGGCTGTGAGCTGCAATGATGCCCTTGGACAGCAACACCTCAAGCAA 4031
QY 1629 AATGACAGCCATGGATATCTCAGATTAATTTGTTGACCACTATTATGACCGC 1688
Db 4032 AATGACAGCCATGGATATCTCAGATTAATTTGTTGACCACTATTATGACCGC 4091
QY 1689 CTGAGCAGAGGACCAACAAATTTGGTCAACGCTCCCTCTCTGCGTGGATATGTGCTGAAC 1748
Db 4092 CTGAGCAGAGGACCAACAAATTTGGTCAACGCTCCCTCTCTGCGTGGATATGTGCTGAAC 4151
QY 1749 TGGCTGTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGCTCTTTTAA 1808

Db 4152 TGGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCGTGTCTGTCTTTAAA 4211
QY 1809 ACTGGCATCATTTCCCTGTCTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG 1868
Db 4212 ACTGGCATCATTTCCCTGTCTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAG 4271
QY 1869 CAAGTGCAAGTTCAACAGGATTTTGACACAGCGAGGCTGGGCTCTCTTCTGCATGAT 1928
Db 4272 CAAGTGCAAGTTCAACAGGATTTTGACACAGCGAGGCTGGGCTCTCTTCTGCATGAT 4331
QY 1929 TCTATCCAAATTTCAACAGCAGTGTGGGTGAAGTTGCACTTTGGGGCAGTAAACATTGAG 1988
Db 4332 TCTATCCAAATTTCAACAGCAGTGTGGGTGAAGTTGCACTTTGGGGCAGTAAACATTGAG 4391
QY 1989 CCAAGTGTCCGGA 2001
Db 4392 CCAAGTGTCCGGA 4404

RESULT 10
US-10-149-736-40
; Sequence 40, Application US/10149736
; Publication No. US20030216332A1
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; APPLICANT: Harper, Scott Q.
; TITLE OF INVENTION: Mini-Dystraphin Nucleic Acids and Peptide Sequences
; FILE REFERENCE: UM-06968
; CURRENT APPLICATION NUMBER: US/10/149,736
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: PCT/US01/31126
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,848
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 5339
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-149-736-40

Query Match 62.7%; Score 1254; DB 13; Length 5339;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1621; Conservative 0; Mismatches 380; Indels 79; Gaps 6;

QY 1 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 60
Db 1199 GGCAGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACAGCTTTAGAGAA 1258
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAACGAGAGAGATTTCTAAT 120
Db 1259 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGAACGAGAGAGATTTCTAAT 1318
QY 121 GATGTGGAAGTGTGAAAGACAGTTTTCATCTCATGAGGGGTACATGATGATTTGACA 180
Db 1319 GATGTGGAAGTGTGAAAGACAGTTTTCATCTCATGAGGGGTACATGATGATTTGACA 1378
QY 181 GCCCATCAGGCGGGTGGTAAATATTTCTCAANTTGGGAAGTGAAGTGGAGACAGGA 240
Db 1379 GCCCATCAGGCGGGTGGTAAATATTTCTCAANTTGGGAAGTGAAGTGGAGACAGGA 1438
QY 241 AAATTATCAGAAGATGAAGAACTGAAGTACAGGACAGATGAATCTCTAAATTTCAAGA 300
Db 1439 AAATTATCAGAAGATGAAGAACTGAAGTACAGGACAGATGAATCTCTAAATTTCAAGA 1498
QY 301 TGGGAATCCCTCAGGGTAGCTAGCATGGAAAAAACAAGCAATTTACATAGATTTTAACTG 360
Db 1499 TGGGAATCCCTCAGGGTAGCTAGCATGGAAAAAACAAGCAATTTACATAGATTTTAACTG 1558
QY 361 GATCTCCGAATC-GAAACTGAAGAGTTGAATGACTGCTAACAACAGAGAAAGA 419

Db 1559 CAACAGTTCCCTCGACCTGGAAAGTTCTTCCCTGGCTTACAGAGGCTGAACAACT 1618
QY 420 ACAAGAAATCGAGGAAGACCTCTTGGACCTGATCTTGAAGACCTAAAGCGCAAGTA 479
Db 1619 GCCAATGCTTACAGGATGCTACCCGTAAGGAAGGCTCTAGAGACTCCNAGGGAGTA 1678
QY 480 CAACACATAAGGTGCTTCAAGAGATCTAGAACAGAAAGTCAGGGTCAATTTCTTC 539
Db 1679 AAGAGCTGATGAACAAATAGGCAAGACCTCCAAAGGTGAATTTGAAGCTCAACAGATGTT 1738
QY 540 ACTCATGTTGGTGGTAG-----TTGATGAATCTAGTGAGATCAGCAACT 587
Db 1739 TATCAACACTGGATGAAGAACAGCCAAATAATCTGAGATCCCTGGAAAGTTCCGATGAT 1798
QY 588 GCTGCTTTGGAAGAACAACTTAAG-----TATGGAGATGATGGGCAAAACATCTGPAGA 644
Db 1799 GCAGTCTCTGTACAAAGACGTTGGATAACATGAATTTCAAGTGGAGTGAATTCGGAAA 1858
QY 645 TGGACAGAAACCGCTGGTTCTTTTACAGACATCCTTCTCAATGGCAACGCTTACT 704
Db 1859 AAGTCTCTCAACATTAGTCCCATTTGGAAGCCAGTTCGACCAAGTGAAGCGTCTGCAC 1918
QY 705 GAAGAACAGTGCCTTTTATGTCATGCTTTTCAAGAAAAGAGATGAGTGAACAAAGATT 764
Db 1919 CTTTCTCTGCAGGAACCTCTGCTGCTACAGCTGAAGATGATGAATTAAGCCGCAG 1978
QY 765 CACACAACTGCTTTAAGATCAAAATGAATGTATCAAGTCTTCAAAACTGGCGGTT 824
Db 1979 GCACCTATTGGAGCGACTTTTCCAGCAGTTCCAGAGCAGAACGATGTACATAGGGCCTTC 2038
QY 825 TTAAGAGCGGATCTAGAAAAGAAAAGCAATCCATGGGCAAACTGTA-----871
Db 2039 AAGAGGNAATTGAATACTAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGATA 2098
QY 872 -TTCACTCAAAAGATCTTCTTTCAACACTGAAGATAAGTCAAGTCAAGCCAGAACGCG 930
Db 2099 TTTCTGACAGAGCAGCTTTTGGAGGACTAGAGAAACTCTACAGGAGCCAGAGAGCTG 2158
QY 931 AAGCATGCTGA-----TAACTTTGCCGCTGTGGG 963
Db 2159 CTTCTGAGGAGAGGCCAGATGTCACTGGGCTTCTAGAAAGAGGCTGAGGAGTCC 2218
QY 964 ATAATTTAGTCCAAAACCTTGAA-----AAGACTACAGCACAG 1001
Db 2219 AATACTGAGTGGGAAAATTTGAACCTGCACCTCGCTGACTGGCAGAGAAAATAGATGAG 2278
QY 1002 ACCCTTGAAGACTCCAGGAACCTTCAAGAGGCGACCGATGAGCTGGAACCTCAAGCTGCGC 1061
Db 2279 ACCCTTGAAGACTCCAGGAACCTTCAAGAGGCGCACCGATGAGCTGGAACCTCAAGCTGCGC 2338
QY 1062 CAAGCTGAGGTGATCAAGGGATCTGGCAGCCCGTGGCGGATCTCCATTGACTCTCTC 1121
Db 2339 CAGCTGAGGTGATCAAGGATCTTGGAGCCCTGGAGCCGATCTCTCATGACTCTCTC 2398
QY 1122 CAAGATCACTCGAAGAAAGTCAAGGCACTTGGAGGAGAAATTCGCTCTGAAAGAGAAC 1181
Db 2399 CAAGATCACTCGAAGAAAGTCAAGGCACTTGGAGGAGAAATTCGCTCTGAAAGAGAAC 2458
QY 1182 GTGAGCCACGTCATGATGCTTCTGCGCAGCTTACCACTTTGGCATTGAGCTCTCACCG 1241
Db 2459 GTGAGCCACGTCATGATGCTTCTGCGCAGCTTACCACTTTGGCATTGAGCTCTCACCG 2518
QY 1242 TATAACCTCAGCACTCTGGAAGACCTTGAACACCAAGATGGAAGCTTCTGCAAGTGGCGCTC 1301
Db 2519 TATAACCTCAGCACTCTGGAAGACCTTGAACACCAAGATGGAAGCTTCTGCAAGTGGCGCTC 2578
QY 1302 GAGACCGAGTCAAGCAGCTGATGAAGCCCAAGGCACTTTGGTCAGGATCTCAGCAC 1361
Db 2579 GAGACCGAGTCAAGCAGCTGATGAAGCCCAAGGCACTTTGGTCAGGATCTCAGCAC 2638
QY 1362 TTTCTTCCAGCTGTCTCAGGGTCCCTGGAGAGGCCATCTCGCCAAACAAGTGGCC 1421

Db 2639 TTTCTTCCACGCTGTGTCCAGGGTCCCTGGGAGAGAGCCATCTCCGCAAAACAAGTGCC 2698
QY 1422 TACTATATCAACACGAGACTCAAAACAATCTGCTGGGACCATCCCAAAATGACAGAGCTC 1481
Db 2699 TACTATATCAACACGAGACTCAAAACAATCTGCTGGGACCATCCCAAAATGACAGAGCTC 2758
QY 1482 TACCACTCTTTAGCTGACCTGAATTAATGTCAAGATCTCAGCTTATAGGACTGCCATGAAA 1541
Db 2759 TACCACTCTTTAGCTGACCTGAATTAATGTCAAGATCTCAGCTTATAGGACTGCCATGAAA 2818
QY 1542 CTCCGAAGACTCAGAGGCCCTTTGCTTGGATCTCTTGGAGCTGTCAAGCTCAGCTCATGTAT 1601
Db 2819 CTCCGAAGACTCAGAGGCCCTTTGCTTGGATCTCTTGGAGCTGTCAAGCTCAGCTCATGTAT 2878
QY 1602 GCTCTGGACAGCAACAACCTCAAGCAAAATGACAGCCCATGGATATCTTCGAGATTAT 1661
Db 2879 GCTCTGGACAGCAACAACCTCAAGCAAAATGACAGCCCATGGATATCTTCGAGATTAT 2938
QY 1662 AATCTTTGACCACTATTATTGACCGCTGGAGCAGAGCAACAATTTGGTCAACGTC 1721
Db 2939 AATCTTTGACCACTATTATTGACCGCTGGAGCAGAGCAACAATTTGGTCAACGTC 2998
QY 1722 CTTCTCTGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACA 1781
Db 2999 CTTCTCTGCTGGATATGTCTGAACTGGCTGCTGAATGTTTATGATACGGGACGAACA 3058
QY 1782 GGGAGATCCGCTGCTCTTTTAAACTGGCATCATTCCTGTGTAAAGCATTGTG 1841
Db 3059 GGGAGATCCGCTGCTCTTTTAAACTGGCATCATTCCTGTGTAAAGCATTGTG 3118
QY 1842 GAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTGTGACCAG 1901
Db 3119 GAAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACACAGGATTTGTGACCAG 3178
QY 1902 CGCAGCTGGGCTCTCTCTGATGATTTCTTCAAAATTCAAAGACAGTTGGGTGAAGTT 1961
Db 3179 CGCAGCTGGGCTCTCTCTGATGATTTCTTCAAAATTCAAAGACAGTTGGGTGAAGTT 3238
QY 1962 GCATCTTTGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 2001
Db 3239 GCATCTTTGGGGCAGTAACATTGAGCCAAAGTGTCCGGA 3278

RESULT 11

US-09-845-416-10
; Sequence 10, Application US/09845416
; Publication No. US2003017312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-10

Query Match 62.5%; Score 1251.2; DB 13; Length 3531;
Best Local Similarity 77.8%; Pred No 0;
Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;
QY 1 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGACCGTTATCAAAACAGCTTTAGAGAA 60
Db 1000 GGCAGTTCATTGATGAGAGTGAAGTAAACCTGACCGTTATCAACAGCTTTAGAGAA 1059
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACACATTCGACAGCAGGAGAGATTTCATAAT 120

1060	Db	GTATTATCGTGGCTTCTTCTGCTGAGGACACATTGCAAGCAACAAGGAGAGATTTCTTAAT	1119
121	QY	GATGTGGAAGTGGTGAAGACCAAGTTTTCATCTCATGAGGGTACATGATGGATTTGACA	180
1120	Db	GATGTGGAAGTGGTGAAGACCAAGTTTTCATCTCATGAGGGTACATGATGGATTTGACA	1179
181	QY	GCCATCAGGGCCGGTGGTAAATATCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGA	240
1180	Db	GCCATCAGGGCCGGTGGTAAATATCTACAATTTGGGAAGTAAGCTGATTTGGAACAGGA	1239
241	QY	AAATTATCAGACATCAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA	300
1240	Db	AAATTATCAGACATCAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA	1299
301	QY	TGGGAATGTCCTCAGGCTAGCTAGCATTTGGAATAACAAGCAATTTACATAGATTTTAATG	360
1300	Db	TGGGAATGTCCTCAGGCTAGCTAGCATTTGGAATAACAAGCAATTTACATAGACTCATAGA	1359
361	QY	GATCTCCAGNA-----TCGAACCTGAAGAGTTGAATGACTTGGCTACAAAAACAGAA	413
1360	Db	TTACTCGAACAGTTCCCTCGGACCTGGAAAAGTTTCTTGCCCTGGCTTACAGAAAGCTGAA	1419
414	QY	GAAAGAACCAAGGAAAATGGAGGAAGACCTCTTTGGACCTGATCTTGAAGACCTTAAACGC	473
1420	Db	ACAACTGCCAATGCTCTACAGGATGCTACCGTAAAGAAAGGCTCCTAGAAGACTCCAAG	1479
474	QY	CAAGTACACACATTAAGGTGCTTCAAGAAGATCTAGAACAGAAACAAGTCAGGTCAT	533
1480	Db	GGAGTAAAGAGCTGATGAACAATGCAAGACCTCCAAAGGTGAATTAAGCTCACACA	1539
534	QY	TCCTCACTCACATGTGTGGTGTAGTTGATGA-----ATCTAGTCGAGATCAC	581
1540	Db	GATGTTTATCACAACTGGATGAAGAACAGCAAAAAAATCCTGAGATCCCTGGAGGTTCC	1599
582	QY	GCAACTGCTGCTTTGGAAGAACAACTTAAG-----TATGGGAGATCGATGGGCAAAATC	638
1600	Db	GATGATGCACTCCTGTTTACAAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAAC	1659
639	QY	TGTAGATGGACAGAACCGCTCGGTTCTTTTACAGACATCCTTCTCAATCGCAACGT	698
1660	Db	CGGAAAAGTCTCTCAACATTAGTCCCATTTGGAGCCAGTTCTGACCAAGTGGAGCGT	1719
699	QY	CTTACTGAAGAACAGTGCCTTTTATGTGCAATGCTTTACAAAAAGAGATGCAGTGAAC	758
1720	Db	CTGCACCTTCTCTGAGGAACCTCTGGTGGCTACAGCTGAAAGATGATGAATTAAGC	1779
759	QY	AAGATTACACAACTGGCTTTAAGATCAAAATGMAATGTTATCAAGTCTTCAAAACTG	818
1780	Db	CGGACGGACCTATTTGAGCGCACTTTCCAGCAGTTCCAGAGCAGAACGATGTACATAGG	1839
819	QY	GCGCTTTTAAAGCGGATCTAGAAAAAGAAAAGCAATCCATGGGCAAACTGTA-----	871
1840	Db	GCCTTCAAGAGGAAATTGAAAACCTAAGAACCCTGTAATCATGAGTACTCTTGAGACTGTA	1899
872	QY	-----TTCACTCAAAACAAGATCTTCTTTCAACACTGAAGAAATAGTCAAGTCAACAG	924
1900	Db	CGAATATTTCTGACAGACGACCTTTTGGAAAGGACTAGAGAACTCTACCAAGAGCCGAG	1959
925	QY	AGCGGAAGCATGSCCTGGA-----TAACTTTGCCCGCGT	957
1960	Db	GAGTGCCTCTGAGGAGAGAGCCGAGAAATGTCACCTGGCTTCTACGAAGACGGCTGAG	2019
958	QY	GTTGGGATAATTAGTCCAAAAACTTGAA-----AGAGTACA	995
2020	Db	GAGTCAATACTAGTGGGAAAAATTGAACTGTCACTCCGCTGACTGCGCAGAGAAAAATA	2079
996	QY	GCAAGACCTTTGAAAGACTCCAGAACTTCAAGAGCCACGGATGAGCTGACCTCAAG	1055
2080	Db	GATGAGACCTTTGAAGACTCCAGAACTTCAAGAGCCACGGATGAGCTGACCTCAAG	2139
1056	QY	CTGCGCCAAGCTGAGGTGATCAAGGGATCTCTGGCAGCCGCTGGCGCATCTCTCAATTGAC	1115

2140	CTGGCCAAAGCTGAGTGTGATCAAGGGATCCTGGCAGCCGCTGGGCGATCTCTCTCATTTGAC	2199
1116	TCTCTCAAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATGCGCCTCTTGAAA	1175
2200	TCTCTCAAGATCACTCTCGAGAAAGTCAAGGCACCTTCGAGGAGAAATGCGCCTCTTGAAA	2259
1176	GAGAACGTGAGCACAAGTCAATGACCTTGTCTGCGACGCTTACACATTTTGGGCATTTCAGCTC	1235
2260	GAGAACGTGAGCACAAGTCAATGACCTTGTCTGCGACGCTTACACATTTTGGGCATTTCAGCTC	2319
1236	TCACCGTATAAACCTCAGCACTCTCGGAAGACCTTGAAACACAGATTGAAGCTTCTGCAAGGTG	1295
2320	TCACCGTATAAACCTCAGCACTCTCGGAAGACCTTGAAACACAGATTGAAGCTTCTGCAAGGTG	2379
1296	GCGGTGAGACCGAGTCAAGGACGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCT	1355
2380	GCGGTGAGACCGAGTCAAGGACGCTGCATGAAGCCACAGGGACCTTTGGTCCAGCATCT	2439
1356	CAGCACCTTTCTTTCCACGCTGTCTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAA	1415
2440	CAGCACCTTTCTTTCCACGCTGTCTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAAACAAA	2499
1416	GTGCGCTACTATATCAACACAGAGACTCAAAACACTTGTCTGGAGCCATCCCAAAATGACA	1475
2500	GTGCGCTACTATATCAACACAGAGACTCAAAACACTTGTCTGGAGCCATCCCAAAATGACA	2559
1476	GAGCTCTACCAAGTCTTTAGCTGACCTCGAATTAATGTCCAGATTTCTCAGCTTATAGGACTGCC	1535
2560	GAGCTCTACCAAGTCTTTAGCTGACCTCGAATTAATGTCCAGATTTCTCAGCTTATAGGACTGCC	2619
1536	ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTGCAGCTGCA	1595
2620	ATGAAACTCCGAAGACTGCAGAAAGGCCCTTTGCTTGGATCTCTTTGAGCCTGTGCAGCTGCA	2679
1596	TGTGATGCTTTGACAGCAGCACACCTTCAGCAAAATGACAGAGCCATGGATATCCCTGTCAG	1655
2680	TGTGATGCTTTGACAGCAGCACACCTTCAGCAAAATGACAGAGCCATGGATATCCCTGTCAG	2739
1656	ATTATTAATTTGTTTGACCACTATTTATGACCGCTCGAGCAAGAGCACAAATTTGGTCT	1715
2740	ATTATTAATTTGTTTGACCACTATTTATGACCGCTCGAGCAAGAGCACAAATTTGGTCT	2799
1716	AAGTCCCTCTCTGCGTGGATATGTCTGAACTGGCTGCTGAATCTTTATGATACCGGA	1775
2800	AAGTCCCTCTCTGCGTGGATATGTCTGAACTGGCTGCTGAATCTTTATGATACCGGA	2859
1776	CGAACAGGAGGATCCGCTGCTCTCTTTTAAAACTGGGCATCATTTCCCTGTGTAAAGCA	1835
2860	CGAACAGGAGGATCCGCTGCTCTCTTTTAAAACTGGGCATCATTTCCCTGTGTAAAGCA	2919
1836	CATTGTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGT	1895
2920	CATTGTGAAGACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGATTTTGT	2979
1896	GACCAGCGCAGGCTGGGCCCTCTTCTGCAATGATTTCTATCCAAATTCCAAGACAGTTTGGGT	1955
2980	GACCAGCGCAGGCTGGGCCCTCTTCTGCAATGATTTCTATCCAAATTCCAAGACAGTTTGGGT	3039
1956	GAACTTGATCTCTTTGGGGGCACTAACTTTGAGCCCAAGTGTCCGGA	2001
3040	GAACTTGATCTCTTTGGGGGCACTAACTTTGAGCCCAAGTGTCCGGA	3085

RESULT 12
US-09-845-416-30
; Sequence 30, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845.416

;; CURRENT FILING DATE: 2001-04-30
;; PRIOR APPLICATION NUMBER: 60/200,777
;; PRIOR FILING DATE: 2000-04-28
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 30
;; LENGTH: 4498
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-845-416-30

Query Match 62.5%; Score 1251.2; DB 13; Length 4498;
Best Local Similarity 77.8%; Pred No. 0;
Matches 1623; Conservative 0; Mismatches 378; Indels 85; Gaps 6;

QY 1 GCAGTTCATTGATGGAGAGTGAATTAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 60
DB 1757 GGCAGTTCATTGATGGAGAGTGAATTAACCTGGACCGTTATCAAAACAGCTTTAGAAGAA 1816
QY 61 GTATTATCGTGGCTTCTTCTGCTGAGGACATTTGCAAGCACAGGAGAGATTTCTAAT 120
DB 1817 GTATTATCGTGGCTTCTTCTGCTGAGGACATTTGCAAGCACAGGAGAGATTTCTAAT 1876
QY 121 GATGTGGAAGTGTGAAAGACAGTTTCACTCATGAGGGGTACATGGAATTTGACA 180
DB 1877 GATGTGGAAGTGTGAAAGACAGTTTCACTCATGAGGGGTACATGGAATTTGACA 1936
QY 181 GCCATCAGGGCCGGTGTGTAATTTCTACAAATTTGGGAAGTACGTTGGAACAGGA 240
DB 1937 GCCCATCAGGGCCGGTGTGTAATTTCTACAAATTTGGGAAGTACGTTGGAACAGGA 1996
QY 241 AAATTATCAGAAGATGAGAAACTCAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA 300
DB 1997 AAATTATCAGAAGATGAGAAACTCAAGTACAAGAGCAGATGAATCTCTAAATTTCAAGA 2056
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAATG 360
DB 2057 TGGGAATGCTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAATG 2116
QY 361 GATCTCCAGAA-----TCGAACTGAAGAGTGAATGATGCTGTAAGACCTTAAACGC 413
DB 2117 TTACTGCAACAGTTCCCGCTGGACCTGGAAGAGTTTCTTGCTGGCTTACAGAGCTGAA 2176
QY 414 GAAAGAACAGAAATGGAGGAAGAGCTCTTGGACCTGATCTTGAAGACCTTAAACGC 473
DB 2177 ACAATGCTCAATGCTCTACAGATGCTACCGTGAAGGAAGGCTCTTGAAGACTCCAG 2236
QY 474 CAAATGACAAACATAAGTGTCTTCAAGAGATCTAGAACAGAACAGTCAAGTCAAT 533
DB 2237 GGAATGAAAGAGCTGATGAAACAAATGGCAAGACCTCCAAAGGTGAAATTTGAAGCTCACACA 2296
QY 534 TCTCTCACTCAATGGTGTGTAGTTGATGA-----ATCTAGTGGAGATCAC 581
DB 2297 GATGTTTATCAACCTGATGAAACAGCCAAATAATCTTGAGATCTCTGGAAGGTTCC 2356
QY 582 GCAATGCTGTGTTGGAAGAACAACTTAAG---TATTGGAGATCGATGGGCAAAATC 638
DB 2357 GATGATGCACTCTGTTACAAAGAGCTTTGGATAACATGAATCAAGTGGAGTGAACCT 2416
QY 639 TGTAGATGCAAGAGACCGCTGGTCTTTTACAGACATCTCTCTCAATGCAACGT 698
DB 2417 CGGAAAAGTCTCTCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCACTGGAAGCGT 2476
QY 699 CTTACTGAAGAACAGTGGCTTTTGTAGTCATGGCTTTTTCAGAAAAGAGATGCAAGTGAAC 758
DB 2477 CTGCACCTTTCTGTCAGGAACCTCTGCTGGTGGCTACAGCTGGAAGATGATGAATTAAGC 2536
QY 759 AAGATTACAAACATGGCTTTTAAGATCAAAATGAATGTTATCAAGTCTTCAAAACCTG 818
DB 2537 CGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTTCAAGAGCAGAACAGTGTACATAGG 2596
QY 819 GCCGTTTTTAAAGCGGACTTAGAAAAGAAAAGCAATCCATGGGCAAACTGTA----- 871

DB 2597 GCCTTCAAGAGGGAATTGAAAACATAAGAACCTGTAAATCATGAGTACTCTTGAGACTGTA 2656
QY 872 -----TTCACTCAAAACAAGATCTTCTTCAACACTGAAGATAAGTCAGTGACCCAGA 924
DB 2657 CGAATATTTCTGACAGAGAGCGCTTTTGAAGAGCTAGAGAAACTCTACAGGAGCCAGA 2716
QY 925 AGACGGAAGCATGGCTGGA-----TAACTTGCCCCGT 957
DB 2717 GAGCTGCTCTCTGAGGAGAGAGCCAGAAATGTCACTCGGCTTCTACGAAAGAGGCTGAG 2776
QY 958 GTTGGGATAATTTAGTCCAAAACATTGAA-----AAGAGTACA 995
DB 2777 GAGTCAATCTAGTGGGAAATTTGAACCTGCACTCCGCTGACTGGGAGAGAAAATA 2836
QY 996 GCACAGACCTTTGAAAGACTCCAGGAACCTCAAGAGGCCACCGATGAGTGGACCTCAAG 1055
DB 2837 GATGAGACCTTTGAAAGACTCCAGGAACCTCAAGAGGCCACCGATGAGTGGACCTCAAG 2896
QY 1056 CTGCGCCAGAGCTGAGTGATCAAGGGATCCTGGACCGCTGGGCGATCTCTCATTTGAC 1115
DB 2897 CTGCGCCAGAGCTGAGTGATCAAGGGATCCTGGACCGCTGGGCGATCTCTCATTTGAC 2956
QY 1116 TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTTGAGGAGAAATTCGGCTCTGAAA 1175
DB 2957 TCTCTCAAGATCACCTCGAGAAAGTCAAGGCACTTTGAGGAGAAATTCGGCTCTGAAA 3016
QY 1176 GAGAACTGAGCAGCTCAATGACCTTGTCTGCGAGCTTACCACCTTTGGGCAATCAGCTC 1235
DB 3017 GAGAACTGAGCAGCTCAATGACCTTGTCTGCGAGCTTACCACCTTTGGGCAATCAGCTC 3076
QY 1236 TCACCTGATAAAGCTCAGCACTCTGGAAGACCTGAAACAGATGGAAGCTTCTGCAAGTG 1295
DB 3077 TCACCTGATAAAGCTCAGCACTCTGGAAGACCTGAAACAGATGGAAGCTTCTGCAAGTG 3136
QY 1296 GCGCTCGAGAACCGAGTCAAGCAGCTGCAATGAAGCCCAAGGCACTTTGGTCCAGCATCT 1355
DB 3137 GCGCTCGAGAACCGAGTCAAGCAGCTGCAATGAAGCCCAAGGCACTTTGGTCCAGCATCT 3196
QY 1356 CAGCACTTTCTTCCAGCTCTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAA 1415
DB 3197 CAGCACTTTCTTCCAGCTCTCCAGGGTCCCTGGGAGAGGCCATCTCGCCAAACAAA 3256
QY 1416 GTGCCCTACTATATCAACCAAGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA 1475
DB 3257 GTGCCCTACTATATCAACCAAGAGACTCAAAACAACTTGTGGGACCATCCCAAAATGACA 3316
QY 1476 GAGCTTACAGTCTTTAGCTGACCTGAATATGTGATGCTCAGATTCTCAGCTTATAGGACTGC 1535
DB 3317 GAGCTTACAGTCTTTAGCTGACCTGAATATGTGATGCTCAGATTCTCAGCTTATAGGACTGC 3376
QY 1536 ATGMAACTCCGAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGGGCTGTGAGCTGCA 1595
DB 3377 ATGMAACTCCGAAGACTGCAAGAGCCCTTTGCTTGGATCTCTTGGGCTGTGAGCTGCA 3436
QY 1596 TGTGATGCTTGGACAGACAACTCAAGCAAAATGACAGCCCATGGATATCTGAG 1655
DB 3437 TGTGATGCTTGGACAGACAACTCAAGCAAAATGACAGCCCATGGATATCTGAG 3496
QY 1656 ATTATTAAATTTGTTGACCACTATTATGACCGCTGGAGAGAGGACCAACAAATTTGCTC 1715
DB 3497 ATTATTAAATTTGTTGACCACTATTATGACCGCTGGAGAGAGGACCAACAAATTTGCTC 3556
QY 1716 AACGTCCTCTCTGCGTGGATGATGTCTGAACTGGCTGCTGAAATGTTTATGATACGGGA 1775
DB 3557 AACGTCCTCTCTGCGTGGATGATGTCTGAACTGGCTGCTGAAATGTTTATGATACGGGA 3616
QY 1776 CGAAGAGGAGATCCGTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 1835
DB 3617 CGAAGAGGAGATCCGTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGCA 3676
QY 1836 CATTTGGAGAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTGT 1895
DB 3677 CATTTGGAGAGCAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGT 3736

QY 1896 GACCAGCGAGCTGGGCTCTTCTGTCATGATTCATCCAAATCCAAAGACAGTTGGGT 1955
DB 3737 GACCAGCGAGCTGGGCTCTTCTGTCATGATTCATCCAAATCCAAAGACAGTTGGGT 3796
QY 1956 GAAGTTGCATCTTTGGGGGAGTAACTAGGCAAGTGTCCGGA 2001
DB 3797 GAAGTTGCATCTTTGGGGGAGTAACTAGGCAAGTGTCCGGA 3842

RESULT 13
US-09-845-416-2
; Sequence 2, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-2

Query Match 62.2%; Score 1245; DB 13; Length 4182;
Best Local Similarity 73.1%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 736; Gaps 2;

QY 1 GGCAGTTCATTGATGGAGAGTGAAGTAACTGACCGTTATCAACAGCTTTAGAGAA 60
DB 1000 GGCAGTTCATTGATGGAGAGTGAAGTAACTGACCGTTATCAACAGCTTTAGAGAA 1059
QY 61 GTATTATCGTGGCTTTCTTCTGCTGAGACACATGTCGAAGCACAAGGAGATTTCTAAT 120
DB 1060 GTATTATCGTGGCTTTCTTCTGCTGAGACACATGTCGAAGCACAAGGAGATTTCTAAT 1119
QY 121 GATGTGGAAGTGTGAAGACCGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
DB 1120 GATGTGGAAGTGTGAAGACCGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 1179
QY 181 GCCATCAGGCGCGGTTGGTAAATTTCTACAATTGGGAAGTAACTGATGGAAACAGGA 240
DB 1180 GCCATCAGGCGCGGTTGGTAAATTTCTACAATTGGGAAGTAACTGATGGAAACAGGA 1239
QY 241 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGA 300
DB 1240 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCCTAAATTCAGA 1299
QY 301 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAACAAAGCAATTTACATAGATTTTAAATG 360
DB 1300 TGGGAATGCCCTCAGGCTAGCTAGCATGGAAACAAAGCAATTTACATAGATTTTAAATG 1359
QY 361 GATCTCCAGATTC-GAACTGAAAGAGTTGAATCACTGGCTTAACAAAACAGAGAAAGA 419
DB 1360 GATCTCCAGATTCAGAACTGAAAGAGTTGAATCACTGGCTTAACAAAACAGAGAAAGA 1419
QY 420 ACAAGGAAATGGAGGAGAGCCCTCTGGACCTCATCTTGAAGACCTAAACCCCAAGTA 479
DB 1420 ACAAGGAAATGGAGGAGAGCCCTCTGGACCTCATCTTGAAGACCTAAACCCCAAGTA 1479
QY 480 CAACAACATAAGGTGCTCTCAAGAGATCTAGAACAGAACAGTCAAGGTCAATTTCTTC 539
DB 1480 CAACAACATAAGGTGCTCTCAAGAGATCTAGAACAGAACAGTCAAGGTCAATTTCTTC 1539
QY 540 ACTCACATGCTGGTGTAGTGTGATGAATCTAGTGGAGATCAACGCACTGCTGTTGGAA 599

DB 1540 ACTCACATGCTGGTGTAGTGTGATGAATCTAGTGGAGATCAACGCAACTGCTGCTTTGGAA 1599
QY 600 GAACAACCTAAAGTATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC 659
DB 1600 GAACAACCTAAAGTATTTGGGAGATCGATGGGCAAAACATCTGTAGATGGACAGAACCGC 1659
QY 660 TGGGTTCTTTTACAAGACATCTCTCAAAATGGCAACGCTTCTACTGAAAGAACAGTGCCTT 719
DB 1660 TGGGTTCTTTTACAAGACATCTCTCAAAATGGCAACGCTTCTACTGAAAGAACAGTGCCTT 1719
QY 720 TTTAGTCATGCTTTTCAGAAAAAGAGATGCAGTGAACAGATTCACACAACCTGCTTT 779
DB 1720 TTTAGTCATGCTTTTCAGAAAAAGAGATGCAGTGAACAGATTCACACAACCTGCTTT 1779
QY 780 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTA 839
DB 1780 AAAGATCAAAATGAAATGTTATCAAGTCTTCAAAACTGGCCGTTTAAAGCGGATCTA 1839
QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACA 899
DB 1840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAAACAGATCTTCTTTCAACA 1899
QY 900 CTGAAGATAAGTCAAGTGAACCCAGAGACGGAAGCATGCTGGATAAATTTGCCCGGTGT 959
DB 1900 CTGAAGATAAGTCAAGTGAACCCAGAGACGGAAGCATGCTGGATAAATTTGCCCGGTGT 1959
QY 960 TGGGATAATTTAGTCCAAAACTTGAAAAGAGTACAGCAC 999
DB 1960 TGGGATAATTTAGTCCAAAACTTGAAAAGAGTACAGCACAGACTCATAGATTACTGCAA 2019
QY 1000 999
DB 2020 CAGTTCGCCCTGGACCTGGAAAAAGTTTCTTGCTGGCTTACAGAAAGCTGAAACAACTGCC 2079
QY 1000 999
DB 2080 AATGTCTACAGATGCTACCCGTAAGGAAGGCTCTTAGAAGACTCCCAAGGGAGTAAAA 2139
QY 1000 999
DB 2140 GAGCTGATGAAACAAATGGCAAGACCTCCAAAGGTGAAATTTGAAGCTCACACAGATGTTAT 2199
QY 1000 999
DB 2200 CACAACCTGGATGAAAAACAGCCAAAAAATCTTGAGATCCCTGGGAAGGTTCCGATGATGCA 2259
QY 1000 999
DB 2260 GTCTCTTACAAAGACGTTTGGATAACATGAATCTCAAGTGGAGTGAATCTCGGAAAAAG 2319
QY 1000 999
DB 2320 TCTCTCAACATTAGTCCCATTTTGGAAAGCCAGTTCTGACCAGTGGAAAGCGTCTGCACCTT 2379
QY 1000 999
DB 2380 TCTCTGCAGGAATCTTCTGTGTGGCTACAGCTGAAGAGATGATGAATTAAGCGGAGSCA 2439
QY 1000 999
DB 2440 CCTATTGGAGGCGACTTTCCAGCAGTTTCAGAAAGCAGAACGATGTACATAGGCGCTTCAAG 2499
QY 1000 999
DB 2500 AGGGAATTTGAAAACTTAAAGAACCTGTATCATGAGTACTCTTGAGACTGTACGAATATTT 2559
QY 1000 999
DB 2560 CTGACAGAGCAGCCTTTTGGAAAGGACTAGAGAAAATCTTACCAGGAGGCCAGAGAGCTGCT 2619
QY 1000 999
DB 2620 CCTGAGGAGAGAGCCCAAGATGTCACTCGGCTTCTACGAAGAGAGGCTGAGGAGGTCAAT 2679

```
QY 1000 -----AGACC 1004
Db 2680 ACTGAGTGGAAAAATTGAACCTGCACCTCCGCTGACTGGCAGAGAAAAATAGATGAGACC 2739
QY 1005 CTTGAAGACTCCAGAACTTCAAGAGGCCACCGGATGAGCTGAGCTCAAGCTGCGCAA 1064
Db 2740 CTTGAAGACTCCAGAACTTCAAGAGGCCACCGGATGAGCTGAGCTCAAGCTGCGCAA 2799
QY 1065 GCTGAGGTGATCAAGGATCTTGGCAGCCGCTGGCGGATCTCTCATTTGACTCTCTCAA 1124
Db 2800 GCTGAGGTGATCAAGGATCTTGGCAGCCGCTGGCGGATCTCTCATTTGACTCTCTCAA 2859
QY 1125 GATCACTCGAGAAAGTCAAGCACTTCGAGAGAAATTTGGCCTCTCAAGAGAACGTTG 1184
Db 2860 GATCACTCGAGAAAGTCAAGCACTTCGAGAGAAATTTGGCCTCTCAAGAGAACGTTG 2919
QY 1185 AGCCACGTCAATGACCTTGTCTGCCAGCTTTACCACTTTGGGCATTCAGCTCTCACCGTAT 1244
Db 2920 AGCCACGTCAATGACCTTGTCTGCCAGCTTTACCACTTTGGGCATTCAGCTCTCACCGTAT 2979
QY 1245 AACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTTCGAGGTGGCGTGAG 1304
Db 2980 AACCTCAGCACTCTGGAAGACCTGAACACAGATGGAAGCTTTCGAGGTGGCGTGAG 3039
QY 1305 GACCGAGTCAGGACGTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGACCTTT 1364
Db 3040 GACCGAGTCAGGACGTGATGAAGCCACAGGAGCTTTGGTCCAGCATCTCAGACCTTT 3099
QY 1365 CTTTCCAGCTGTCTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTAC 1424
Db 3100 CTTTCCAGCTGTCTCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTCCCTAC 3159
QY 1425 TATATCAACCCAGGACCTCAACACCTTCTGGGACCATCCCAATTCACAGACCTTAC 1484
Db 3160 TATATCAACCCAGGACCTCAACACCTTCTGGGACCATCCCAATTCACAGACCTTAC 3219
QY 1485 CAGCTTTAGCTGACCTGAATAATGTCAAGTCTCAGCTTATAGGACTGCCATGAAACTC 1544
Db 3220 CAGCTTTAGCTGACCTGAATAATGTCAAGTCTCAGCTTATAGGACTGCCATGAAACTC 3279
QY 1545 CGAGACTGCGAGAGGCCCTTTGCTTGATCTTTGAGCCTGTGAGTGCATGATGATGCC 1604
Db 3280 CGAGACTGCGAGAGGCCCTTTGCTTGATCTTTGAGCCTGTGAGTGCATGATGATGCC 3339
QY 1605 TTGGACACGACCAACCTCAAGCBAATGACAGCCCATGGATATCTCTGCAGATTATTAAT 1664
Db 3340 TTGGACACGACCAACCTCAAGCBAATGACAGCCCATGGATATCTCTGCAGATTATTAAT 3399
QY 1665 TGTGTGACCACTATTATGACCGCTGGAGCAAGACACAAATTTGGTCAAGCTCCCT 1724
Db 3400 TGTGTGACCACTATTATGACCGCTGGAGCAAGACACAAATTTGGTCAAGCTCCCT 3459
QY 1725 CTCTGCGTGGATATGCTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1784
Db 3460 CTCTGCGTGGATATGCTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 3519
QY 1785 AGGATCCCTGTCTCTTTTAAACTGGCATCAATTTCCCTGTGTAAGACACATTTGGAA 1844
Db 3520 AGGATCCCTGTCTCTTTTAAACTGGCATCAATTTCCCTGTGTAAGACACATTTGGAA 3579
QY 1845 GACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGTGACCAAGCC 1904
Db 3580 GACAAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTGTGACCAAGCC 3639
QY 1905 AGGCTGGCCCTCTCTGATGATTTCTATCCAAATTCACAGAGTGGGTGAAGTTGCA 1964
Db 3640 AGGCTGGCCCTCTCTGATGATTTCTATCCAAATTCACAGAGTGGGTGAAGTTGCA 3699
QY 1965 TCCTTTGGGGCAGTAACATTTGAGCCCAAGTGTCCGGA 2001
Db 3700 TCCTTTGGGGCAGTAACATTTGAGCCCAAGTGTCCGGA 3736
```

RESULT 14

```
US-09-845-416-27
; Sequence 27, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 50/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 27
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-27
```

Query Match 62.2%; Score 1245; DB 13; Length 5149;

Best Local Similarity 73.1%; Pred. No. 0;

Matches 2001; Conservative 0; Mismatches 0; Indels 736; Gaps 2;

```
QY 1 GCAGGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA 60
Db 1757 GCAGGTTCAATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAACACAGCTTTAGAGAA 1816
QY 61 GTATTATCTGGCTTCTTCTGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAAT 120
Db 1817 GTATTATCTGGCTTCTTCTGCTGAGGACACATTGCAAGCAACAGGAGAGATTTCTAAT 1876
QY 121 GATGTGAAGTGGTGAAGACCAAGCTTCACTCATGAGGGGTACATGATGATTTGACA 180
Db 1877 GATGTGAAGTGGTGAAGACCAAGCTTCACTCATGAGGGGTACATGATGATTTGACA 1936
QY 181 GCCCATCAGGGCGGGTGGTAAATTTCTACAAATGGGAGGTAAAGCTGATTCGACAGGA 240
Db 1937 GCCCATCAGGGCGGGTGGTAAATTTCTACAAATGGGAGGTAAAGCTGATTCGACAGGA 1996
QY 241 AAATATATCAGAAGATGAAGAACTGAAAGTCAAGAGCAGATGAATCTCTCTAAATCAAGA 300
Db 1997 AAATATATCAGAAGATGAAGAACTGAAAGTCAAGAGCAGATGAATCTCTCTAAATCAAGA 2056
QY 301 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAAACAAGCAATTTACATAGATTTTAATG 360
Db 2057 TGGGAATGCTCAGGGTAGCTAGCATGGAAAAAACAAGCAATTTACATAGATTTTAATG 2116
QY 361 GATCTCCAGAACT-GAAACTGAACAGTTGAATGATGGCTAAACAAAAACAGAAAGAA 419
Db 2117 GATCTCCAGAACTGAAGAACTGNAAGAGTTGAATGATGGCTAAACAAAAACAGAAAGAA 2176
QY 420 ACAAGAAAAATGGAGAAAGAGCTCTTGGACCTGATCTTGAAGACCTAAACGCGCAAGTA 479
Db 2177 ACAAGAAAAATGGAGAAAGAGCTCTTGGACCTGATCTTGAAGACCTAAACGCGCAAGTA 2236
QY 480 CAACACATTAAGTGTTCACAGAGATCTAGAACAGAACAGTCAAGTCAAGTCAATTTCTC 539
Db 2237 CAACACATTAAGTGTTCACAGAGATCTAGAACAGAACAGTCAAGTCAAGTCAATTTCTC 2296
QY 540 ACTCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 599
Db 2297 ACTCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2356
QY 600 GAACAACTTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 659
Db 2357 GAACAACTTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2416
QY 660 TGGGTTCTTTTACAGACATCTCTCTCAATGGCAACGCTCTTACTGAGAAACAGTGCCTT 719
Db 2417 TGGGTTCTTTTACAGACATCTCTCTCAATGGCAACGCTCTTACTGAGAAACAGTGCCTT 2476
```



```

/ FILE REFERENCE: UM-06968
/ CURRENT APPLICATION NUMBER: US/10/149,736
/ CURRENT FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: PCT/US01/31126
/ PRIOR FILING DATE: 2001-10-04
/ PRIOR APPLICATION NUMBER: 60/238,848
/ PRIOR FILING DATE: 2000-10-06
/ NUMBER OF SEQ ID NOS: 96
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 41
/ LENGTH: 5462
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic
/ US-10-149-7336-41

```

Query Match 56.6%; Score 1132.2; DB 13; Length 5462;
Best Local Similarity 73.9%; Pred. No. 0;
Matches 1628; Conservative 0; Mismatches 373; Indels 202; Gaps 7;

QY	1	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTCGACCGGTTATCAAAACAGCTTTAGAGAA	60
Db	1199	GGCAGTTCATTGATGGAGAGTGAAGTAAACCTCGACCGGTTATCAAAACAGCTTTAGAGAA	1258
QY	61	GTATTATCGTGGCTTCCTTTCTGCTGAGGACACATTCGACAGCAACAGGAGAGATTTCTTAAT	120
Db	1259	GTATTATCGTGGCTTCCTTTCTGCTGAGGACACATTCGACAGCAACAGGAGAGATTTCTTAAT	1318
QY	121	GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGTAGGGGTACATGATGAGATTTCACA	180
Db	1319	GATGTGGAAGTGGTGAAGACCAAGTTTCATCTCATGTAGGGGTACATGATGAGATTTCACA	1378
QY	181	GCCCATCAGGCCCGGGTTGGTAATTTCACAATGGGAAAGTAAGCTGATTGGAACAGGA	240
Db	1379	GCCCATCAGGCCCGGGTTGGTAATTTCACAATGGGAAAGTAAGCTGATTGGAACAGGA	1438
QY	241	AAATTATCAGAGAGATGAAGAACTGAAGTACAGAGCAGATCAATCTCCTTAATTTCAAGA	300
Db	1439	AAATTATCAGAGAGATGAAGAACTGAAGTACAGAGCAGATCAATCTCCTTAATTTCAAGA	1498
QY	301	TGGGAATGCCCTCAGGCTAGCTAGCATGGAAAAACAAGCAAATTACATAGAGATTTTAATG	360
Db	1499	TGGGAATGCCCTCAGGCTAGCTAGCATGGAAAAACAAGCAAATTACATAGAGATTTTAATG	1558
QY	361	-----GATCTCCAGNATCGAA-----	360
Db	1559	ACCACATTGGAGGCGCTCCTCTACTCAGACTGTTTACTCTGGTGACACAACCTGTGGTTACT	1618
QY	361	-----GATCTCCAGNATCGAA-----	377
Db	1619	AAGGAACCTGCCATCTCCAACTAGAAATGCCATCTTCCTGATGTTGGAGCATAGATTA	1678
QY	378	-----CTGAAGAGAGTTGNAATGCACTGGCTAACCAAAAACAGAGAA	416
Db	1679	CTGCAACAGTTTCCCCCTGGACCTGGAAAGTTTCTTCCCTGGCTTACAGAGCTGAAACA	1738
QY	417	AGAACAGGAAAAATGGAGGAGAGCCCTTTGGACCTGATCTTGAAGACCTTAAACGCCAA	476
Db	1739	ACTGCCAATGTCTTACAGGATGTCACCCGTGAAGAAAGGCTCCTAGAAGACTCCAAGGGA	1798
QY	477	GTACACAACATATAGGTGCTTCAAGAGATCTAGACAGACAGTCAAGGTCAATCTCT	536
Db	1799	GTAAAGAGCTGTATGAACAAATGGCAAGACCTCCAAGGTGAATTTGAAGCTCACACAGAT	1858
QY	537	CTCACTACATGGTGGTGGTAGTTGATGA-----ATCTAGTGGAGATCAGCGCA	584
Db	1859	GTTTATCACACCTGGATGAACAGCCAAAAAATCCTGCAGATCCCTGGAAGGTTCCGAT	1918
QY	585	ACTGCTGCTTTGGAGAACAACTTAAAG---TATTGGGAGATCGATGGGCAAAAATCTGT	641
Db	1919	GATGCAGTCTCTGTACAAAGACGTTTGGATAACATGAATTCAGGTGGAGTGAACCTCGG	1978

```
Db      |||||||ATTAATGTTGACCACTATTATGACCGCTGGACGACACACACATTTGGTCAAC 3118
Qy      |||||||GTCCCTCTGCGTGGATATGTCGTCTGAAGTGGCTGCTGAAATGTTTATGATACGGGACGA 1778
Db      |||||||GTCCCTCTGCGTGGATATGTCGTCTGAAGTGGCTGCTGAAATGTTTATGATACGGGACGA 3178
Qy      |||||||ACAGGGAGGATCCGTGTCCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACAT 1838
Db      |||||||ACAGGGAGGATCCGTGTCCTGCTCTTTTAAACTGGCATCATTTCCCTGTGTAAGCACAT 3238
Qy      |||||||TTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGAC 1898
Db      |||||||TTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGAC 3298
Qy      |||||||CAGCGAGGCTGGGCTCCTCTCTGCAATGATTCATCCAAATCCAAAGACAGTTGGTGAA 1958
Db      |||||||CAGCGAGGCTGGGCTCCTCTCTGCAATGATTCATCCAAATCCAAAGACAGTTGGTGAA 3358
Qy      |||||||GTTGCATCCTTTGGGGCAGTAAACATTGAGCCCAAGTGTCCGGA 2001
Db      |||||||GTTGCATCCTTTGGGGCAGTAAACATTGAGCCCAAGTGTCCGGA 3401
```

Search completed: February 2, 2004, 17:40:10
Job time : 697.123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2004, 11:04:35 ; Search time 3752.09 seconds
(without alignments)
12961.661 Million cell updates/sec

Title: US-09-845-416-14_COPY_1000_3000

Perfect score: 2001
Sequence: 1 ggcagtcattgatgagag.....cattgagccaagtgtccgga 2001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	49.5	3870	11	BC036103 Homo sapi
2	776.6	38.8	3056	11	AK044536 Mus muscu
3	585.2	29.2	728	14	CB228986 AGENCOURT
4	569	28.4	824	9	AL556247 AL556247

5	567.4	28.4	1098	13	EX365572
6	566.4	28.3	801	14	CB991394
7	558	27.9	620	13	BQ640063
8	532.2	26.6	770	10	BG19710
9	511.4	25.6	1298	11	AK087829
10	511.4	25.6	2135	11	AK013510
11	511.4	25.6	4437	11	AK036936
12	511	25.5	797	14	CB960722
13	509.8	25.5	1384	11	AK075809
14	494.6	24.7	652	10	BB629984
15	491.6	24.6	704	10	BB610411
16	467.2	23.3	854	9	AI196693
17	460.6	23.0	3753	11	AK081426
18	455.8	22.8	591	2	HSN083521
19	455.8	22.8	717	14	CB527785
20	439	21.9	578	2	HSN075761
21	436.6	21.8	665	14	BY742604
22	422.8	21.1	898	10	BF182065
23	417.6	20.9	843	14	CA988247
24	410.4	20.5	599	10	BB666688
25	407.2	20.3	554	14	CB613696
26	404.4	20.2	493	14	CA888041
27	402.6	20.1	1490	11	BC009242
28	397.4	19.9	495	14	CA894775
29	397	19.8	595	14	CB177816
30	392.2	19.6	650	14	BY714491
31	385.2	19.3	684	9	AL641565
32	383.4	19.2	483	14	CA893902
33	381.8	19.1	423	9	AA460476
34	380.4	19.0	633	9	AL796733
35	375.6	18.8	794	13	BU424348
36	375.4	18.8	514	9	AL871560
37	365.4	18.3	655	14	CA558919
38	349.6	17.5	546	4	BX518597
39	345.8	17.3	3051	11	BC036095
40	337.6	16.9	541	9	AL894729
41	337	16.8	665	10	BE201973
42	332.6	16.6	2334	11	BC011062
43	330	16.5	608	9	AL672616
44	329	16.4	532	10	BE334408
45	325.6	16.3	532	12	BU036615

ALIGNMENTS

RESULT 1
BC036103
LOCUS Homo sapiens, clone IMAGE:5274415, mRNA.
DEFINITION BC036103
ACCESSION BC036103
VERSION BC036103.1 GI:23271310
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3870)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

BC036103 3870 bp mRNA linear HTC 04-MAR-2003
Homo sapiens, clone IMAGE:5274415, mRNA.

BC036103.1 GI:23271310
HTC.
Homo sapiens (human)

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3870)

Strausberg, R.

Direct Submission

Submitted (31-JUL-2002)

National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., De Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Cariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashina, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851

5
 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 3056)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akanishi, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muranatsu, M. and Hayashizaki, Y.

Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscl.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Retina RNA was provided by Dr. Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers

source
 1. 3056
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:A930019P21"
 /db_xref="taxon:10090"
 /clone="A930019P21"
 /tissue_type="retina"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"

misc_feature
 1. 3056
 /note="dyestrophin, muscular dystrophy (MGD|MGF:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)"

BASE COUNT 1072 a 567 c 607 g 810 t

ORIGIN

Query Match 38.8%; Score 776.6; DB 11; Length 3056;
 Best Local Similarity 86.5%; Pred. No. 2.9e-153;

Matches 868; Conservative 0; Mismatches 134; Indels 1; Gaps 1;
 QY 1 GGCAGTTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTATTCAACACAGCTTTAGAGAA 60
 Db 1044 GACAGTTTCATTGATGGAGAGTGAAGTAAACCTGGAGTAGTATCCAACTGCTTTAGAGAA 1103
 QY 61 GTATATTCGTGGCTTCCTTCTGTGAGGACACATGTCAGACCAAGAGAGATTTCTAAT 120
 Db 1104 GTACTTTTCATGGCTTCCTTCTGTGAGGATACATGTCGAGCAAGAGAGATTTCAAT 1163
 QY 121 GATGTGAAGTGTGGAAGACCACTGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 180
 Db 1164 GATTTGAAGTGTGGAAGACCACTGTTTCATCTCATGAGGGGTACATGATGGATTTGACA 1223
 QY 181 GCCCATCAGGCGCGGTGGTGAATATTTCTCAATTTGGAAAGTAAAGTGTGGAACAGGA 240
 Db 1224 TCTCATCAAGGACTTGTGGTGAATTTCTAGTTAGGAAGTCAACTAGTTGGAAGAGG 1283
 QY 241 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 300
 Db 1284 AAATTATCAGAGATGAAGAACTGAAGTACAGAGCAGATGAATCTCTTAATTTCAAGA 1343
 QY 301 TGGGAATGCTTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 360
 Db 1344 TGGGAATGCTTCAGGGTAGCTAGCATGGAAGAAACAAAGCAATTTACATAGAGTTTAAATG 1403
 QY 361 GATCTCCAGATTC-GAAATCTGAAGAGTGAATGATCTGGCTTACAAAGAGAGAGAAAGA 419
 Db 1404 GATCTCCAGATTCAGAAATTTAAAGAACTAGATGACTGGTTTAAAGAAAGTGAAGAGAGA 1463
 QY 420 ACAAGGAAATGGAGGAGAGAGCTCTTTGGACCTGATCTTCAAGACCTTAAACCGCAAGTA 479
 Db 1464 ACTAAGAAATGGAGGAGAGAGCTCTTTGGACCTGATCTTCAAGACCTTAAATGCCAAGTA 1523
 QY 480 CAACACATTAAGTGTCTTCAAGAGATCTTAGAACAAAGCAAGTCAAGGTCAATTTCTCTC 539
 Db 1524 CAACACATTAAGTGTCTTCAAGAGATCTTAGAACAAAGCAAGTCAAGGTCAATTTCTCTC 1583
 QY 540 ACTCACATGTTGGTGTAGTGTGATGATCTTAGTGGAGATCACGCAACTGCTGCTTTGGA 599
 Db 1584 ACTCACATGTTGGTGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1643
 QY 600 GAACAACTTAAGTGTATTTGGAGAGATCGATGGGCAACATCTGTAGATGCAGAGAGACCGC 659
 Db 1644 GAACAACTTAAGTGTATTTGGAGAGATCGATGGGCAACATCTGTAGATGCAGAGAGACCGC 1703
 QY 660 TGGTTCTTTTACAGACATCTCTTCTCAATTTGGCAAGTCTTACTGAGAACAGTGGCTT 719
 Db 1704 TGGATTGTTTACAGATATTTCTTAAATTTGGCAGCAATTTTACTGAGAACAGTGGCTT 1763
 QY 720 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAAGTGAACAAAGATTTCACAACTGGCTT 779
 Db 1764 TTTAGTGCATGGCTTTTCAGAAAAAGAGATGCAAGTGAACAAAGATTTCACAACTGGCTT 1823
 QY 780 AAAGATCAAAATGAATGTTTATCAAGTCTTCAAAAGTGGCGCTTTTAAAGCGGATCTA 839
 Db 1824 AAAGATCAAAATGAATGTTTATCAAGTCTTCAAAAGTGGCGCTTTTAAAGCGGATCTA 1883
 QY 840 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCACTCAAAAGATCTTCTTTCAACA 899
 Db 1884 GAAAGAAAAAGCAATCCATGGGCAAACTGTATTCTCACTCAAAAGATCTTCTTTCAACA 1943
 QY 900 CTGAAGAAATGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 959
 Db 1944 CTGAAGAAATGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2003
 QY 960 TGGGATTAATTTAGTTCAAAAAATTGAAAGAGTACAGCAGAGA 1002
 Db 2004 TGGGACAAATTTAACCCAAAAAATTGAAAGAGTTCAGCAGAAA 2046

RESULT 3
 CB228986

LOCUS CB228986 728 bp mRNA linear EST 10-FEB-2003
DEFINITION AGENCOURT 1149247 NICHDRH.Ov1 Macaca mulatta cDNA clone
IMAGE:6884820 5', mRNA sequence.
ACCESSION CB228986
VERSION CB228986.1 GI:28280564
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 728)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rga@nci.nih.gov
Tissue Procurement: Dr. Eliot Spindel
CDNA Library Preparation: CLONTECH
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM3135 row: e column: 11
High quality sequence stop: 583.
FEATURES
Location/Qualifiers
1..728
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884820"
/tissue_type="Ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDRH.Ov1"
/notes="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHDRH.Ov1
BASE COUNT 256 a 127 c 162 g 182 t 1 others
ORIGIN
Query Match 29.28; Score 585.2; DB 14; Length 728;
Best Local Similarity 96.98; Pred. No. 5.8e-113;
Matches 618; Conservative 0; Mismatches 18; Indels 2; Gaps 2;
195 GGTGGTGAATATCTACAAATTGGGAAGTAACTGCTGATGGACACAGGAAAATTATCAGAAGA 254
1 GGTGGTGAATATCTACAAATTGGGAAGTAACTGCTGATGGACACAGGAAAATTATCAGAAGA 60
255 TGAAGAACTGAAGTACAGACAGATGAATCTCTTAATTCAGATGGGATGCTCAG 314
61 TGAAGAACTGAAGTACAGACAGATGAATCTCTTAATTCAGATGGGATGCTCAG 120
315 GGTAGCTAGCATGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATC- 373
121 GGTAGCTAGCATGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAAATCA 180
374 GAAACTGAAGAGTTGAATGATGGCTTAACAAAACAGAAAGAACACAGGAAATGGA 433
181 GAAACTGAAGAGTTGAATGATGGCTTAACAAAACAGAAAGAACACAGGAAATGGA 240
434 GGAAGAGCCCTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACAAACAAATAGGT 493
241 GAAGAAACCCCTTGGACCTGATCTTGAAGACCTTAAACGCAAGTACAAACAAATAGGT 300
494 GCTTCAAGAGATCTAGAAACAAAGCAAGTCAGGGTCAATTTCTCACTCACATGGTGGT 553
301 GCTTCAAGAGATCTAGAAACAAAGCAAGTCAGGGTCAATTTCTCACTCACATGGTGGT 360
554 GGTAGTTGATGAATCTAGTGGAGATCAGCACTGCTGCTTTGGAAGAACAACTTAAGGT 613

361 GGTAGTTGATGAATCTAGTGGAGATCAGCAACTGCTGCTTTTGAAGAGCAACTTAAGGT 420
614 ATTGGGAGATCGATGGCAAAACATCTGTAGATGACAGAGACCGCTGGGTCTTTTAC 673
421 ATTGGGAGATCGATGGCAAAACATCTGTAGATGACAGAGACCGCTGGGTCTTTTAC 480
674 AGACATCTTCTCAATGGCAACGCTTCTTACTAGAACAGTGCCTTTTATAGTCATGGCT 733
481 AGACATCTTCTCAATGGCAACGCTTCTTACTAGAACAGTGCCTTTTATAGTCATGGCT 540
734 TTCAGAAAAGAGATGACGTGACAGAGATTCACAACT-GGCTTTAAAGATCAAAATG 792
541 TTCAGAAAAGAGATGCTGTGAACAGATTCACAACTGGGCTTTAAGATCAAAATG 600
793 AATGTTTATCAAGTCTTCAAAAACACTGCCCTTTTAAAA 830
601 AATGTTTATCAAGTCTTCAAAAACACTGCCCTGTATGAA 638
RESULT 4
AL556247
LOCUS AL556247 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION "cDNA clone CS0DK001VB17 5-PRIME, mRNA sequence."
ACCESSION AL556247
VERSION AL556247.2 GI:31278051
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12898746.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence clone 1955.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DK001CA09Qp1.
FEATURES
Location/Qualifiers
1..824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001VB17"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="HELA"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 211 a 210 c 206 g 197 t
ORIGIN
Query Match 28.4%; Score 569; DB 9; Length 824;
Best Local Similarity 100.0%; Pred. No. 1.5e-109;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1433 CCAGGAGACTCAAAACACTTCTGCGGACCATCCAAAATGACAGAGCTTACAGTCTTT 1492
195 CCAGGAGACTCAAAACACTTCTGCGGACCATCCAAAATGACAGAGCTTACAGTCTTT 254
1493 AGCTGACCTGAATATGTGAGTATCTCAGCTTATAGAGTCCCAATGAACTCCGAGACT 1552
255 AGCTGACCTGAATATGTGAGTATCTCAGCTTATAGAGTCCCAATGAACTCCGAGACT 314

```
QY 1553 GCAGAGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGTCATGCTTGGACCA 1612
Db 315 GCAGAGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGTCATGCTTGGACCA 374
QY 1613 GCACAACTCAAGCAAAATGACCAAGCCATGGATATCTCTCAGATTAATTTGTTGAC 1672
Db 375 GCACAACTCAAGCAAAATGACCAAGCCATGGATATCTCTCAGATTAATTTGTTGAC 434
QY 1673 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGTCAGATTAATTTGTTGAC 1732
Db 435 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGTCAGATTAATTTGTTGAC 494
QY 1733 GGATATGTCGTAACCTGGCTGCTGTAATTTATGATACGGGAGCAACAGGAGGATCCG 1792
Db 495 GGATATGTCGTAACCTGGCTGCTGTAATTTATGATACGGGAGCAACAGGAGGATCCG 554
QY 1793 TGTCTGTCTTTAAACTGTCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTA 1952
Db 555 TGTCTGTCTTTAAACTGTCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTA 614
QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCGAGCGAGGCTGGG 1912
Db 615 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCGAGCGAGGCTGGG 674
QY 1913 CCTCTCTGTCATGATCTATCCAAATCCAAAGACAGTTGGGTGAAGTTGCATCTTTGG 1972
Db 675 CCTCTCTGTCATGATCTATCCAAATCCAAAGACAGTTGGGTGAAGTTGCATCTTTGG 734
QY 1973 GGGCAGTAACATTGAGCAAGTGTCCGGA 2001
Db 735 GGGCAGTAACATTGAGCAAGTGTCCGGA 763
```

```
RESULT 5
BX365572 1098 bp mRNA linear EST 05-MAY-2003
LOCUS BX365572 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK001YB21 5-PRIME, mRNA sequence.
ACCESSION BX365572
VERSION BX365572.1 GI:30366927
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1098)
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1955.r
Contact: Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSIAK001ZB11QPL
Location/Qualifiers
1..1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YB21"
/cell_line="HELA"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="Test strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 286 a 271 c 268 g 262 t 11 others
```

FEATURES

source

ORIGIN

```
Query Match 28.4%; Score 567.4; DB 13; Length 1098;
Best Local Similarity 99.8%; Pred. No. 3.5e-109;
Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1433 CCACGAGACTCAAAACAATTTGCTGGACCATCCCAAATGACAGAGCTCTACCACTCTTT 1492
Db 204 CCACGAGACTCAAAACAATTTGCTGGACCATCCCAAATGACAGAGCTCTACCACTCTTT 263
QY 1493 AGCTGACCTGATATATGTCAGATTTCTAGCTTTATAGGACTCCATGAACTCCGAAGACT 1552
Db 264 AGCTGACCTGATATATGTCAGATTTCTAGCTTTATAGGACTCCATGAACTCCGAAGACT 323
QY 1553 GCAGAAGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGTCATGATGCTTGGACCA 1612
Db 324 GCAGAAGCCCTTTGCTTGGATCTCTTGGAGCTGTGATGTCATGATGCTTGGACCA 383
QY 1613 GCACAACTCAAGCAAAATGACCAAGCCATGGATATCTCTCAGATTAATTTGTTGAC 1672
Db 384 GCACAACTCAAGCAAAATGACCAAGCCATGGATATCTCTCAGATTAATTTGTTGAC 443
QY 1673 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGTCAGATTAATTTGTTGAC 1732
Db 444 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGTCAGATTAATTTGTTGAC 503
QY 1733 GGATATGTCGTAACCTGGCTGCTGTAATTTATGATACGGGAGCAACAGGAGGATCCG 1792
Db 504 GGATATGTCGTAACCTGGCTGCTGTAATTTATGATACGGGAGCAACAGGAGGATCCG 563
QY 1793 TGTCTGTCTTTTAAACTGTCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTA 1852
Db 564 TGTCTGTCTTTTAAACTGTCATCATTTCCCTGTGTAAAGCAATTTGGAAGCAAGTA 623
QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCGAGCGAGGCTGGG 1912
Db 624 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCGAGCGAGGCTGGG 683
QY 1913 CCTCTCTGTCATGATCTATCCAAATCCAAAGACAGTTGGGTGAAGTTGCATCTTTGG 1972
Db 684 CCTCTCTGTCATGATCTATCCAAATCCAAAGACAGTTGGGTGAAGTTGCATCTTTGG 743
QY 1973 GGGCAGTAACATTGAGCAAGTGTCCGGA 2001
Db 744 GGGCAGTAACATTGAGCAAGTGTCCGGA 772
```

RESULT 6

CB991394

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB991394 801 bp mRNA linear EST 01-MAY-2003
AGENCOURT 13627932 NIH_MGC 148 Homo sapiens cDNA clone
IMAGE:30336570 5', mRNA sequence.
CB991394 GI:30285818
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 801)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: NDAM360 row: d column: 19
High quality sequence stop: 621.
Location/Qualifiers
1..801
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30336570"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tcon"
/clone_lib="NIH_MGC_148"
/note="Organ: Placenta; Vector: pBluescriptR; Site: 1: all-XhoI; Site 2: BamHI; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH MGC Library."
204 a 210 c 199 g 188 t

BASE COUNT 204 a 210 c 199 g 188 t

ORIGIN

Query Match 28.3%; Score 566.4; DB 14; Length 801;
Best Local Similarity 99.8%; Pred. No. 5.5e-109;
Matches 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1433 CCACGAGACTCAAAACACTTGTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 1492
Db 140 CCACGAGACTCAAAACACTTGTGGGACCATCCCAAATGACAGAGCTCTACAGTCTTT 199
QY 1493 AGCTGACCTGAATATGTGAGATTCAGCTTATAGGATGCCATGAACTCCGAGACT 1552
Db 200 AGCTGACCTGAATATGTGAGATTCAGCTTATAGGATGCCATGAACTCCGAGACT 259
QY 1553 GCAGAGGCGCTTGTGCTGATCTTGGAGCTGTGAGCTGTGATGATGCTTGACCA 1612
Db 260 GCAGAGGCGCTTGTGCTGATCTTGGAGCTGTGAGCTGTGATGATGCTTGACCA 319
QY 1613 GCACAACTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTTAATGTTTGC 1672
Db 320 GCACAACTCAAGCAAAATGACAGCCCATGATATCTTGCAGATTTAATGTTTGC 379
QY 1673 CACTATTTATGACCGCTGGAGCAAGACACAACTTTGTCACAGTCCCTCTCGGT 1732
Db 380 CACTATTTATGACCGCTGGAGCAAGACACAACTTTGTCACAGTCCCTCTCGGT 439
QY 1733 GGATATGTGCTGAACCTGGCTGCTGAATGTTATGATACGGACCAACAGGAGGATCCG 1792
Db 440 GGATATGTGCTGAACCTGGCTGCTGAATGTTATGATACGGACCAACAGGAGGATCCG 499
QY 1793 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACATTTGGAGACAGTA 1852
Db 500 TGTCTGTCTTTTAAACTGGCATCATTTCCCTGTGTAAAGACATTTGGAGACAGTA 559
QY 1853 CAGATACCTTTTCAAGCAAGTGCAGTTCACAGGATTTTGTGACCGAGCGTGGG 1912
Db 560 CAGATACCTTTTCAAGCAAGTGCAGTTCACAGGATTTTGTGACCGAGCGTGGG 619
QY 1913 CTTCTTCTGCATGATTTATCCAAATTCAGACAGTGGGTGAGTTGCATTCCTTGG 1972
Db 620 CTTCTTCTGCATGATTTATCCAAATTCAGACAGTGGGTGAGTTGCATTCCTTGG 679
QY 1973 GGCAGTAACATTGAGCCAGTGTCCGG 2000
Db 680 GGCAGTAACATTGAGCCAGTGTCCGG 707

RESULT 7
B0640063 620 bp mRNA linear EST 15-JUL-2002
LOCUS B0640063 Human Retina cDNA (Un-normalized, unamplified): hd/he
DEFINITION He23904.y1

Homo sapiens cDNA clone he23904 5', mRNA sequence.
B0640063
VERSION B0640063.1 GI:21764522
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 620)
AUTHORS J.W. Bouffard, G. Smith, D. and Peterson, K.
TITLE Expressed sequence tag analysis of human retina for the NEIBank Project: Rebinding, an abundant, novel retinal cDNA and alternative splicing of other retina-preferred gene transcripts
JOURNAL Mol. Vis. 8 (4), (2002) In press
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 23 row: g column: 04
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="he23904"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retina cDNA (Un-normalized, unamplified): hd/he"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue was dissected from two 80 year old donors with no observed eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel, MD) essentially following the protocols of the Superscript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTCTAGATCGAGCGCCGCC(T)15-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."
165 a 148 c 144 g 163 t

BASE COUNT 165 a 148 c 144 g 163 t

ORIGIN

Query Match 27.9%; Score 558; DB 13; Length 620;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1444 AAACAACCTTCTGGGACCATCCCAAATGACAGCTCTACAGTCTTTAGCTGACCTGA 1503
Db 1 AAACAACCTTCTGGGACCATCCCAAATGACAGCTCTACAGTCTTTAGCTGACCTGA 60
QY 1504 ATATGTGAGATCTCTGAGCTTATAGGACTGCCATGAACTCCGAAGACTCGAAGGCC 1563
Db 61 ATAATGTGAGATCTCTGAGCTTATAGGACTGCCATGAACTCCGAAGACTCGAAGGCC 120
QY 1564 TTTCTTGGATCTCTGAGCTGTGAGCTGTGATGCTTGGACCAACACTCA 1623
Db 121 TTTCTTGGATCTCTGAGCTGTGAGCTGTGATGCTTGGACCAACACTCA 180
QY 1624 AGCAAAATGACCGCCCATGATATCTGAGATTTAATGCTTTGACCACTATTATG 1683
Db 181 AGCAAAATGACCGCCCATGATATCTGAGATTTAATGCTTTGACCACTATTATG 240
QY 1684 ACCGCTGGAGCAAGACACAAATTTGGTCAACGTCCTCTCTGCTGATATGTGTC 1743


```
Query Match      26.6%; Score 532.2; DB 10; Length 770;
Best Local Similarity 98.1%; Pred. No. 8.9e-102;
```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishii, K., Kitsuina, T., Tashiro, H., Itoh, M.,
 3
 20499374
 11042159

Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 Sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staehli,F., Suzuki,R., Tomita,M., Wagner,L., Wastio,T., Sakai,K., Okido,T., Furuno,M., Kono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., De Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,J., Kohtsuki,S. and Hayashizaki,Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11218551

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1298)

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
 Direct Submission

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

SOURCE

FEATURES

SOURCE

SOURCE

SOURCE

SOURCE

/strain="C57BL/6J"
 /db xref="FANTOM DB:E330026B12"
 /db xref="taxon:10090"
 /clone="E330026B12"
 /sex="female"
 /tissue_type="ovary"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="2 days pregnant adult"
 120..1142
 /note="unnamed protein product; dystrophin, muscular dystrophy (MGD|GI:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)
 /putative"
 /codon_start=1
 /protein_id="BAC40019.1"
 /db xref="GI:26352778"
 /translation="MREHLXGHTQTTCWDPKNTLYQSLADNNVFSAYTAMKL RRLQKALCLDLLSLAACDGLDNRKLVKQNDPMDIQINCLTIYDLRQEHNNLVN VPLQCDMLNLLNVDTGRTIVLSFKTIGIISLCKAHLEKDYRLFPQVNASSTGF CDORLGLLHDSIQIPIQGLGEVAFSGSNIEPSVRSFOPANNKPEIEALFLDMWR LEPSQMWLPVLRHVAARAEAKQAKNCKECPILIGFRYSLKHPNDYCSCFFSG RVAKGHKHIFPWEYCTPTTSGEDVRDFAKVLNKKRTKRYFAKHPMGILPQVTLE GDNNET"
 BASE COUNT 343 a 332 c 296 g 327 t
 ORIGIN
 Query Match 25.6%; Score 511.4; DB 11; Length 1298;
 Best Local Similarity 93.7%; Pred. No. 2.3e-97;
 Matches 533; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
 1433 CCACGAGACTCAAAACAACTTGTCTGGGACCATCCAAAATGACAGAGTCTACAGTCTTT 1492
 140 CCACGAGACCCAAACCACTTGTGGGACCAACCCCAAAATGACAGAGTCTACAGTCTTT 199
 1493 AGCTGACCTGAATATGTGAGTCTTGTGAGCTCTTGTGAGCTCTGATGCTGATGCTGAC 1552
 200 AGCTGACCTGAATATGTGAGTCTTGTGAGCTCTTGTGAGCTCTGATGCTGATGCTGAC 259
 1553 GCAGAGGCGCTTGTGCTTGTGATCTTGTGAGCTCTTGTGAGCTCTGATGCTGATGCTGAC 1612
 260 CCAGAGGCGCTTGTGCTTGTGATCTTGTGAGCTCTTGTGAGCTCTGATGCTGATGCTGAC 319
 1613 GCACAACTCAAGCAAAATGACAGCCCATGATGATGATGATGATGATGATGATGATGATGAC 1672
 320 GCACAACTCAAGCAAAATGACAGCCCATGATGATGATGATGATGATGATGATGATGATGAC 379
 1673 CACTATTTATGACCGCTGGAGCAAGACCAACAATTTGTGCTCAACCTCTCTCTGCT 1732
 380 TACAAATTTATGATGCTTGGAGCAAGAGCAACAATCTGGTCAATGCTCTCTCTGCT 439
 1733 GGATATGTCTGAACTGGCTGCTGAAATTTATGATACCGGACGAAACAGGAGGATCCG 1792
 440 GGATATGTCTCAACTGGCTTCTCAATGTTATGATACGCGACGAAACAGGAGGATCCG 499
 1793 TGTCCTGCTTTTAAACTGGCATCTTCCCTGCTGTAAAGCACATTTGGAAGCAAGTA 1852
 500 TGTCCTGCTTTTAAACTGGCATCTTCTCTGTAAAGCACATTTGGAAGCAAGTA 559
 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGACCGCAGCGCTGGG 1912
 560 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACTGGCTTTTGTGACCGCTAGGCTGGG 619
 1913 CTTCTCTTGTGATGATTTTATCCAAATTTCAAGACAGTTGGTGAAGTGCATCTTTGG 1972
 620 TCTTCTTCTGATGATTTCTTATTTCAATCCCAAGACAGTTGGTGAAGTGTCTTCTTTGG 679
 1973 GGGCAGTAACTTGAAGCCCAAGTGTCCGCA 2001
 680 GGGCAGTAACTTGAAGCCCAAGTGTCCGCA 708
 RESULT 10
 AK013510

RESULT 10

AK013510

1733 GGATATGTCTCAACTGGCTGCTGAATGTTTATGATACGGGACCAACAGGAGGATCGG 1792
 Db 444 GGATATGTCTCAACTGGCTGCTGAATGTTTATGATACGGGACCAACAGGAGGATCGG 503

1793 TGCTCTGCTTTTAAACTGGCATCAATCCCTGCTGTAAGACACATTTGGAGACAAAGTA 1852
 Db 504 TGCTCTGCTTTTAAACTGGCATCAATCCCTGCTGTAAGACACATTTGGAGACAAAGTA 563

1853 CAGATACCTTTTCAAGCAAGTGCAGAGTTCAACAGGATTTTGGACCGAGCGGCTGGG 1912
 Db 564 CAGATACCTTTTCAAGCAAGTGCAGAGTTCAACAGGATTTTGGACCGAGCGGCTGGG 623

1913 CCTCTCTGTCATGATTTCTATCCAAATTCACAGACATTTGGGTGAAGTTGTCATCTCTTTGG 1972
 Db 624 TCTCTCTGTCATGATTTCTATCCAAATTCACAGACATTTGGGTGAAGTTGTCATCTCTTTGG 683

1973 GGGCAGTAACATTCAGCAAGTGCAGG 2001
 Db 684 GGGCAGTAACATTCAGCAAGTGCAGG 712

RESULT 11
 AK036936 LOCUS
 DEFINITION Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930028B14 product:dystrophin, muscular dystrophy, full insert sequence.

ACCESSION AK036936
 VERSION 1
 KEYWORDS GI:26085504
 SOURCE HPC; CAP trapper.
 ORGANISM Mus musculus (house mouse)

REFERENCE 1
 Carninci, P. and Hayashizaki, Y. Sugahara, N., Hayatsu, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

2
 20499174
 11042159

3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

4
 20530913
 11076861

5
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schiraldi, L. M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anno, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaeerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

21085560
 11217851

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

6 (bases 1 to 4437)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@ec.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers
 1. 4437
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:9930028B14"
 /db_xref="taxon:10090"
 /clone="9930028B14"
 /sex="female"
 /tissue_type="vagina"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"

misc_feature
 29. 1968
 /note="dystrophin, muscular dystrophy (MGD|MG1:94909, GI|NM_007868, evidence: BLASTN, 100%, match=999)
 putative"
 4417. 4422
 /note="putative"
 4437
 /note="putative"

polyA_signal
 4417. 4422
 /note="putative"
 4437
 /note="putative"

polyA_site
 4437
 /note="putative"

BASE COUNT 1300 a 932 c 885 g 1320 t
 ORIGIN

Query Match 25.6%; Score 511.4; DB 11; Length 4437;
 Best Local Similarity 93.7%; Pred. No. 2.7e-97;

Matches 533; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1433 CCACGAGACTCAAAACACTTCTGGGACCATCCAAATACAGAGCTCTACAGTCTTT 1492

DB 49 CCACGAGACCCAAACACTTCTGGGACCCACCAATACAGAGCTCTACAGTCTTT 108

QY 1493 AGCTGACCTGAATATGTCAGATCTCAGCTTATAGACTGCCATGAAATCCGAGACT 1552

DB 109 AGCTGACCTGAATATGTCAGATCTCAGCTTATAGACTGCCATGAAATCCGAGACT 168

QY 1553 GCAGAAGCCCTTTGCTGGATCTTTGAGCTCTGAGCTGCATGTCATGCTTGGACCA 1612

DB 169 CGAGAAGCCCTTTGCTGGATCTTTGAGCTCTGAGCTGCATGTCATGCTTGGACCA 228

QY 1613 GCACAACCTCAAGCAAAATGACCAAGCCCATGATATCTCGAGATTAATTTGTTGAC 1672

DB 229 GCACAACCTCAAGCAAAATGACCAAGCCCATGATATCTCGAGATTAATTTGTTGAC 288

QY 1673 CACTATTATACCGCTGGAGCAAGACCAACAATTTGGTCAACCTCCCTCTCTCGGT 1732

DB 289 TACAATTTATGATCGTCTGGAGCAAGACCAACAATTTGGTCAACCTCCCTCTCTCGGT 348

QY 1733 GGATATGTCCTGACTGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 1792

DB 349 GGATATGTCCTGACTGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 408

QY 1793 TGCTCTGCTTTTAAACCTGCATCATTTCCCTGTGTAAGACATTTGGAAGACAAGTA 1852

DB 409 TGCTCTGCTTTTAAACCTGCATCATTTCTCTGTGTAAGACACATTTGGAAGACAAGTA 468

QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCAAGCGCAGGCTGGG 1912

DB 469 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTACCAAGCGCAGGCTGGG 528

QY 1913 CCTCCTTCTGCATGATTCATTCACAAATTCACAGAGTGGGTGAAGTTCATCCTTTGG 1972

DB 529 TCTTCTTCTGCATGATTCATTCACAAATTCACAGAGTGGGTGAAGTTCATCCTTTGG 588

QY 1973 GGGCAGTAACATTGAGCAAGTGTCCGA 2001

DB 589 GGGCAGTAACATTGAGCGAGTGTCCGA 617

RESULT 12

CB960722 797 bp mRNA linear EST 29-APR-2003

LOCUS AGENCOURT_13761995 NIH_MGC_147 Homo sapiens cDNA clone

DEFINITION IMAGE:30344479 5', mRNA sequence.

ACCESSION CB960722

VERSION CB960722.1 GI:30216839

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 797)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM380 row: n column: 08
High quality sequence stop: 620.
Location/Qualifiers
1. .797

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30344479"

/tissue_type="Human Placenta"

/lab_host="DH10B Tona"

/clone_lib="NIH_MGC_147"

/note="Organ: placenta; Vector: pBluescriptR; Site.1: all-XhoI; Site.2: BamH; Oligo-dr primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to RCT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: This is a NIH MGC library."

BASE COUNT 199 a 208 c 199 g 191 t

ORIGIN

Query Match 25.5%; Score 511; DB 14; Length 797;

Best Local Similarity 99.8%; Pred. No. 2.7e-97;

Matches 522; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1433 CCACGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 1492

DB 168 CCACGAGACTCAAAACACTTGTGGGACCATCCCAAAATGACAGAGCTCTACAGTCTTT 227

QY 1493 AGCTGACCTGAATATGTCAGATCTCAGCTTATAGACTGCCATGAAATCCGAGACT 1552

DB 228 AGCTGACCTGAATATGTCAGATCTCAGCTTATAGACTGCCATGAAATCCGAGACT 287

QY 1553 GCAGAAGCCCTTTGCTGGATCTCTTGAGCCTGTGTCAGCTGCATGTCATGTCGACCA 1612

DB 288 GCAGAAGCCCTTTGCTGGATCTCTTGAGCCTGTGTCAGCTGCATGTCATGTCGACCA 347

QY 1613 GCACAACCTCAAGCAAAATGACCAAGCTTCATGATATCTTGGAGATTAATTTGTTGAC 1672

DB 348 GCACAACCTCAAGCAAAATGACCAAGCTTCATGATATCTTGGAGATTAATTTGTTGAC 407

QY 1673 CACTATTATGACCGCTGGAGCAAGACCAACAATTTGTCACAGTCCCTCTCTCGGT 1732

DB 408 CACTATTATGACCGCTGGAGCAAGACCAACAATTTGTCACAGTCCCTCTCTCGGT 467

QY 1733 GGATATGTCCTGAACTGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 1792

DB 468 GGATATGTCCTGAACTGCTGCTGAATGTTTATGATACGGGACGAACAGGAGGATCCG 527

QY 1793 TGCTCTGCTTTTAAACCTGCGATCATTTCCCTGTGTAAGACATTTGGAAGACAAGTA 1852

DB 528 TGCTCTGCTTTTAAACCTGCGATCATTTCCCTGTGTAAGACATTTGGAAGACAAGTA 587

QY 1853 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCGCTGGG 1912

DB 588 CAGATACCTTTTCAAGCAAGTGGCAAGTTCACAGGATTTTGTGACCAAGCGCTGGG 647

QY 1913 -CCTCCTTCTGCATGATTCATTCACAAATTCACAGAGTGGG 1954

DB 648 CCTCCTTCTGCATGATTCATTCACAAATTCACAGAGTGGG 690

RESULT 13

AK075809

LOCUS

DEFINITION AK075809 1384 bp mRNA linear HTC 07-DEC-2002

AK075809 Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810074E05 product:dystrophin, muscular dystrophy, full insert sequence.

ACCESSION AK075809

VERSION AK075809.1 GI:26344655

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

- 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
- 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
- 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamanoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Iogawa,Y., Izawa,M., Ohara,E., Watahiki,N., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11078861
- 4
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gotohori,T., Bono,H., Kasukawa,T., Saito,R., Kadoya,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Giesi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaudo,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Cariboldi,M., Guscincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L., Wyshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851

REFERENCE

5
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

- TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 1217851
PUBMED 1217851
- 6
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hasegawa,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Toyama,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toyama,T., Yamanaka,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source
1. 1384
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:1810074E05"
/db_xref="XGI:1908112"
/db_xref="taxon:10090"
/clone="1810074E05"
/sex="male"
/tissue type="pancreas"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 day old"
120. 1142
/note="unnamed protein product; dystrophin, muscular dystrophy (MGD|G1:94909, GB|NM_007868, evidence: BLASTN, 100%, match=999)
putative"
/codon_start=1
/protein_id="BAG35977.1"
/db_xref="G1:26344656"
/db_xref="XGI:94909"
/translation="MREHLKGHETQTTWDPHPMTLYOSLADLNVRFSAYTAMKL RPLQKLDLILSLAACDALQNDQMDIQLINCLTIYDLRQHNLYN RVLCMDLNLNVYDTGTGRIVLSFKTIGISLKHALEDKRYLFPQVASTGF CDQRGLGLLHSDIQIPLQGEVAFSGSNIEPVSFCOFANNKHEIEALFLDWR LEQSSWMLPVLHRAAATAKQAKNICKEPILIGFRYSKHKENYDLCQSCFFRS RVAKGHKQHWIPMVEYCTPTTSGEDVRDPAKLVNKKFRKYPFAKHPMGLPQTVLE GDNNET"

CDS

BASE COUNT 376 a 354 c 303 g 351 t
ORIGIN
Query Match 25.5%; Score 509.8; DB 11; Length 1384;
Best Local Similarity 93.5%; Pred.No. 5.1e-97;
Matches 532; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 1433 CCACGAGACTCAAAACACTTGTCTGGGACCATCCAAAATGACAGAGCTTACCAGTCTTT 1492
DB 140 CCACGAGACCCAAACACTTGTCTGGGACCATCCAAAATGACAGAGCTTACCAGTCTTT 199
QY 1493 AGCTGACCTGATATGTCAGATCTCAGCTTATAGGACTGCGATGAACTCCGAGACT 1552
DB 200 AGCTGACCTGATATGTCAGATCTCAGCTTATAGGACTGCGATGAACTCCGAGACT 259
QY 1553 GCAGAAAGCCCTTGTCTGGATCTCTTGGACCTCTCAGCTGTCATGTGATCCCTTGGACCA 1612
DB 260 CCAGAGAGCCCTTGTCTGGATCTCTTGGACCTCTCAGCTGTCATGTGATCCCTTGGACCA 319
QY 1613 GCACAACTCAAGAAAATGACAGCCCATGATATCTCTGATATCTCTGATATCTCTGATATCT 1672
DB 320 GCACAACTCAAGAAAATGACAGCCCATGATATCTCTGATATCTCTGATATCTCTGATATCT 379
QY 1673 CACTATTATGACCGCTGGAGCAAGAGCAACAATTTGGTCAACGCTCCCTCTCTGCGT 1732
DB 380 TACAATTATGATCTCTGAGCAGAGACACAACTCTGCTCAATCTCTCTCTCTGCTGT 439
QY 1733 GGATATGTGTGTAACCTGGCTGTGTAATGTTTATGATPACGGGACGAAACAGGAGGATCCG 1792
DB 440 GGATATGTGTGTAACCTGGCTGTGTAATGTTTATGATPACGGGACGAAACAGGAGGATCCG 499

